Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report 141	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE APPLICATION OF DISTRIBUTED SYSTEM ARCHITECTURES FOR COMPUTER-BASED INSTRUCTIONAL SYSTEMS IN THE NAVAL EDUCATION AND TRAINING COMMAND		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Barney L. Capehart and Charles L. Morris, Jr.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Training Analysis and Evaluation Group Department of the Navy Orlando, FL 32813		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE February 1983 13. NUMBER OF PAGES 84
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Distributed Processing
Computer Managed Instruction
Computer Based Instruction
Local Area Networks
Data Communications

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report examines the technology of distributed data processing systems in terms of its application to Navy computer based instruction system implementation. It presents the concept of distributed systems, discusses the major components and design considerations, and considers the benefits of applying this technology to individualized instruction for Naval training.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TAEG

TRAINING
ANALYSIS
AND
EVALUATION
GROUP

TECHNICAL REPORT 141

THE APPLICATION OF DISTRIBUTED SYSTEM ARCHITECTURES FOR COMPUTER-BASED INSTRUCTIONAL SYSTEMS IN THE NAVAL EDUCATION AND TRAINING COMMAND

FEBRUARY 1983

FOCUS ON THE TRAINED PERSON

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION IS UNLIMITED.

DEPARTMENT OF THE NAVY · CHIEF OF NAVAL EDUCATION AND TRAINING

TRAINING ANALYSIS AND EVALUATION GROUP
ORLANDO, FLORIDA 32813

Technical Report 141

THE APPLICATION OF DISTRIBUTED SYSTEM ARCHITECTURES FOR COMPUTER-BASED INSTRUCTIONAL SYSTEMS IN THE NAVAL EDUCATION AND TRAINING COMMAND

Barney L. Capehart
Charles L. Morris, Jr.

Training Analysis and Evaluation Group

February 1983

Alfred F. Smode

ALFRED F. SMODE, Ph.D., Director
Training Analysis and Evaluation Group

W. L. Maloy

W. L. MALOY, Ed.D.
Deputy Chief of Naval Education and Training for Educational Development and Research and Development

ACKNOWLEDGMENTS

The authors acknowledge the contributions of the TAEG staff members who were assembled to develop the CNET Computer Based Instruction Plan. During the operational concept planning phase, this group identified a need for a Navy training data processing and communications network which led to the development of this report.

The authors specifically acknowledge the contributions of Mr. Morris Middleton and Dr. M. Michael Zajkowski. Mr. Middleton provided the initial technical direction for the study and made a number of technical contributions relating to system design. Dr. Zajkowski formulated the presentation approach for the introductory sections which served to clarify the relationships among computer based instruction, training efficiency objectives, and distributed processing, and contributed to the preparation of the report.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF ILLUSTRATIONS

LIST OF TABLES

SECTION I

INTRODUCTION

The Chief of Naval Education and Training (CNET) is responsible for the management of over 4,000 courses of instruction with as many as 250,000 enrollees per year. Because of the high cost of providing this instruction, the effectiveness and efficiency of developing and delivering instruction are critical concerns. Increasingly, attention is being focused on computers as one means of improving the effectiveness and efficiency of these functions. Attention is also being directed toward developing methods for effectively integrating and phasing this technology into Navy training programs. Fundamental to the latter concern is the issue of centralized versus distributed computer system architecture. A clear understanding of the logic underlying these alternative approaches is a necessary prerequisite to an informed procurement policy for the efficient use of computer-based training in the Naval Education and Training Command (NAVEDTRACOM).

BACKGROUND

The use of computers to support instruction in the Navy had its roots in the programmed instruction and computer assisted instruction (CAI) work of the 1950s and 1960s. The emergence of this technology combined with the efforts of the Office of Naval Research, Naval Personnel and Training Research Laboratory, and Chief of Naval Air Technical Training to improve the effectiveness and efficiency of training culminated in the development of the Navy Computer Managed Instruction (CMI) system in 1972.

In 1974, CNET and the Chief of Naval Technical Training (CNTECHTRA) adopted CMI as a formal component of the Navy Training System (CNETINST 5260.1, CNTECHTRAINST 5400.7A). The first course to be officially implemented into the CMI system was Aviation Fundamentals at the Naval Air Technical Training Center, Memphis. CNET Decision Memorandum No. 2 (27 April 1976) integrated CMI with the plans for the redesign of courses by the Instructional Program Development Centers. Today, approximately 9,000 students in 9 courses receive their initial technical training via this system.

The major features of the CMI system include centralized evaluation and prescription (management) of self-paced instruction. The instruction, however, is provided off-line at widely separated geographic locations. Several studies (Middleton, Papetti, and Micheli, 1974; Zajkowski, Heidt, Corey, Mew, and Micheli, 1979) have concluded that as training requirements increase in response to changes in the complexity and technology of weapon systems, individualized instruction will of necessity become the instructional strategy of choice and widespread use of computers will be required to support this strategy.

In the decade since the implementation of the CMI system there has been an explosive growth in the use of minicomputers and microcomputers for education and training functions. Initially, applications of this

technology addressed labor intensive management functions. But increasingly, this technology is being used to improve the preparation and delivery of instruction. The Navy Personnel Research and Development Center (NPRDC), Office of Naval Research (ONR), and Naval Training Equipment Center (NTEC) have research programs directed toward this latter purpose. NPRDC has exploratory development projects (6.2) in computer based techniques for training tactical knowledge and for training procedural skills. In advanced development (6.3) NPRDC is examining a Low Cost Microcomputer Training System, Remote Site Training using microprocessors, Advanced Computer Based Training for propulsion engineering, Procedures for Decentralizing Course Design, and Authoring Instructional Materials. NTEC is investigating low cost applications of Emerging Computer Technology and Microcomputer Architecture for Training Systems. ONR efforts are organized around two major themes. The first program area deals with principles of Human and Machine Intelligence for computer based training and job aiding. The second program area deals with learning principles and their application to training. The objective of the program is to build a comprehensive theory of the effects of the structure and content of instructional materials and to use this knowledge to produce guidelines for the design of instruction. The outcomes of these programs are expected to provide significant benefits to military training.

Computer-based instruction activity in the Navy is not restricted to the research environment. A large number of microcomputers are in use to perform routine administrative and word processing functions. In addition, CNET subordinate commands are increasing the use of CBI to directly support schoolhouse training. For example, the Computer Aided Interactive Testing System (CAITS) is being used to test and remediate students at the Airman Apprentice School, Great Lakes. The Electronic Equipment Maintenance Trainer (EEMT), a generic computer based training device, is being integrated into the curriculum at Electronics Technician School at Great Lakes. Computer aided instruction is being implemented and evaluated at Cryptologic Technician Operator "A" School, Corry Station. The effectiveness and efficiency of these and other applications are being carefully analyzed in terms of the ability of this technology to meet Navy training requirements. In view of technological advances (such as described above), it appears that the current Navy CMI might become obsolete in the 1985 to 1995 timeframe. With this awareness, CNET tasked TAEG to conduct a study of the current and projected state of the art in the technology associated with the management and delivery of instruction as a basis for developing future initiatives. This study (Micheli, Morris, and Swope, 1980) identified several important trends and concluded the following:

- DOD will be unable to maintain conventional methods of instruction oriented around stand-up instruction and printed materials in the face of budget austerities.

- Small computers with limited memory will be available to provide test answer analyses and feedback, prompting help, and diagnostics in a course of instruction.

- Costs of microprocessors, memory modules, and components which provide the computer capability for graphic display will continue to decrease for the next decade.

- Computer assisted instruction (CAI) will become more common as computer-based delivery systems become more widely used.

- Intelligent courseware, or intelligent CAI, will increase rapidly.

- Totally distributed instructional delivery systems will be available.

The need to improve training efficiency and the continual advances in training/computer technology compelled CNET to take two important actions related to computer based instruction. First, CNET established the policy (CNETINST 3920.1B) that automation of training functions will be considered whenever there is a reasonable likelihood of increases in training productivity. Second, CNET tasked the TAEG to develop a preliminary plan for the wide-scale application of computer based instruction in the NAVEDTRACOM.[1,2] The first step in implementing the plan was the establishment of the Computer Aided Instruction Working Group (CNET 01C) as a special assistant to the CNET. This group is responsible for implementing and managing a long range plan for integrating CBI into the NAVEDTRACOM. Among the major issues to be resolved by this group is the identification of cost effective CBI prototype configurations which can meet CNET training requirements. A knowledge of the capabilities and shortfalls of centralized and distributed systems is essential to maximal employment of CBI in training.

PURPOSE

This report provides a compilation and integration of information dealing with distributed system architectures for computer-based instructional systems. It is a single source of rapid orientation to the concepts of distributed network architectures, computer resource distribution, computer communication networks, and local area distributed processing in computer-based instruction. It will serve as a guide for NAVEDTRACOM managers in planning, acquiring, integrating, and operating distributed microcomputer systems for training.

ORGANIZATION OF THE REPORT

Distributed computer systems take many forms. However, they all consist of definable subsystem elements and functional concepts. Accordingly, this report is organized in a manner which reflects a generally accepted basic set of elements for distributed system architectures. In addition to this introduction, the report contains six other sections. Section II discusses the concept of distributed systems and compares decentralized and centralized systems. It also discusses potential distributed system applications for Naval training. Section III addresses distributed system network architectures and describes many of the generic network topologies. Section IV provides an overview of resource distribution considerations such as the distribution of storage, communications, and processing. Section V discusses the communications elements of the distributed architectures and addresses topics

[1]CNET ltr N-5 of 3 March 1982.

[2]CNET ltr N-5 of 10 March 1982.

such as transmission mediums, network switching, protocols, data security, and currently used networks. Section VI covers the recently popular topic of local area networks which are geographically limited distributed systems. These local area networks are extremely important since they provide the structure for office automation systems and most likely for CBI systems of the future. Section VII addresses distributed processing applications relating to computer-based instruction and more specifically to examples of CBI within NAVEDTRACOM.

This report organization provides the training system decision maker with information relating to network architectures for computer-based training. It is also of sufficient depth to provide guidelines to the computer systems designer. Since this dual readership has been considered throughout the preparation of this report, it is possible that some sections might be primarily management-directed while others appear excessively technical. If the reader is perplexed by this duality, additional information may be obtained from the many references cited herein.

SECTION II

THE CONCEPT OF DISTRIBUTED SYSTEMS

In the late 1960s the large, mainframe computer provided the heart of most large-scale information systems. This highly centralized approach to information systems was the result of technological advances in the design of central processing units, and the attendant economies of scale which reduced the average cost of computational capability per processing unit. These large, centralized computer systems offered many advantages through the use of proven technology, centralized control of operations, vendor software support, and the availability of turn-key systems for a variety of applications.

Although large scale, centralized computer information systems offer many advantages, they also suffer from a number of problems. A major problem is a lack of user responsiveness. As more and more end users are connected to a centralized system, the amount of time that can be assigned to a single user decreases, and response time becomes intolerably long. A centralized system is also slow to respond to changes in functional requirements because large systems cannot easily be expanded in small increments.

Other disadvantages which reduce the attractiveness of large, centralized systems involve system reliability and availability. A failure in the central processor will usually result in a shutdown of the entire system. If a terminal fails, only one user is affected, but when the main processing unit fails, all users are affected. One solution to the reliability problem, although it is often costly, is to duplicate the critical central processor, thereby reducing the probability of total system failure.

The trend toward use of highly centralized information systems was slowed, and ultimately reversed, by the capability and availability of low cost minicomputers and microcomputers in the mid- to late-1970s. These advances made it possible to put a significant amount of processing power at the end user location. As aptly stated by one practitioner, distributed processing is "putting computer power where the people and problems are." The basic concept of a distributed system is to perform some of the information processing tasks at the user activity location, while allowing other tasks to be communicated to a higher level in the system for processing. The following discussion of distributed systems provides an overview of the technology and a set of guidelines for evaluating alternative distributed system designs.

COMPARING CENTRALIZED AND DECENTRALIZED SYSTEMS

In order to evaluate the potential benefits offered by a distributed system design, it is necessary to compare it to alternative designs, one of which is usually a centralized system. This subsection describes a typical centralized computer system and discusses its advantages and disadvantages so that the contrast with a decentralized distributed system will be clear.

CENTRALIZED SYSTEMS. A schematic diagram of a typical large centralized computer system is shown in figure 1. A large number of video display terminals are connected through communication links to a large, mainframe computer which accomplishes all of the processing and data storage. The remote terminals are strictly input and output devices without any information processing or storage capability. The central computer schedules communications with the terminals, assigns a slice of processing time to each terminal user task, and handles all of the processing and information storage and retrieval functions. This system design is an excellent example of many present centralized information processing systems. As such, it will be the object of discussion in the following paragraphs.

Large centralized computer systems are very expensive, primarily because of the cost of the mainframe computer and the collection of video terminals. The overall system design is inflexible, in that the system can only be expanded in relatively large increments, and, because of the centralized architecture, system reliability and availability are serious concerns.

Many of these problems can be alleviated or eliminated by using a distributed system design. A distributed system is typically very flexible and more responsive and reliable than a centralized system. In addition, system cost is normally less than a comparable centralized system. Distributed system features which bring about these improvements are discussed in the following paragraphs.

DECENTRALIZED SYSTEMS. Before proceeding directly to a description of distributed systems, it is useful to discuss another possible approach to an alternative system design. This approach is essentially that of a totally decentralized system; the very opposite alternative to a fully centralized system. Since microcomputers have become so inexpensive and so capable, some users have simply implemented a system consisting of a collection of individual microcomputers that are not connected together in any way, with stand-alone capability for each unit. The advantages of this design approach are that the total system is relatively inexpensive, single unit failures do not affect other units, and response time is quick because there is no competition for the processor resource.

There are, however, some significant disadvantages of totally decentralized systems. For instance, there is no centralized system control and no automated flow of management information. Additionally, there is no resource sharing which results in a requirement for a large number of peripheral components. System hardware and software configuration control also becomes more difficult, and without some form of central control overall mission accomplishment may be adversely affected.

In some instances, a totally decentralized system might be very appropriate to meet very specific needs. However, most large systems will require the use of centralized control and the flow of management information.

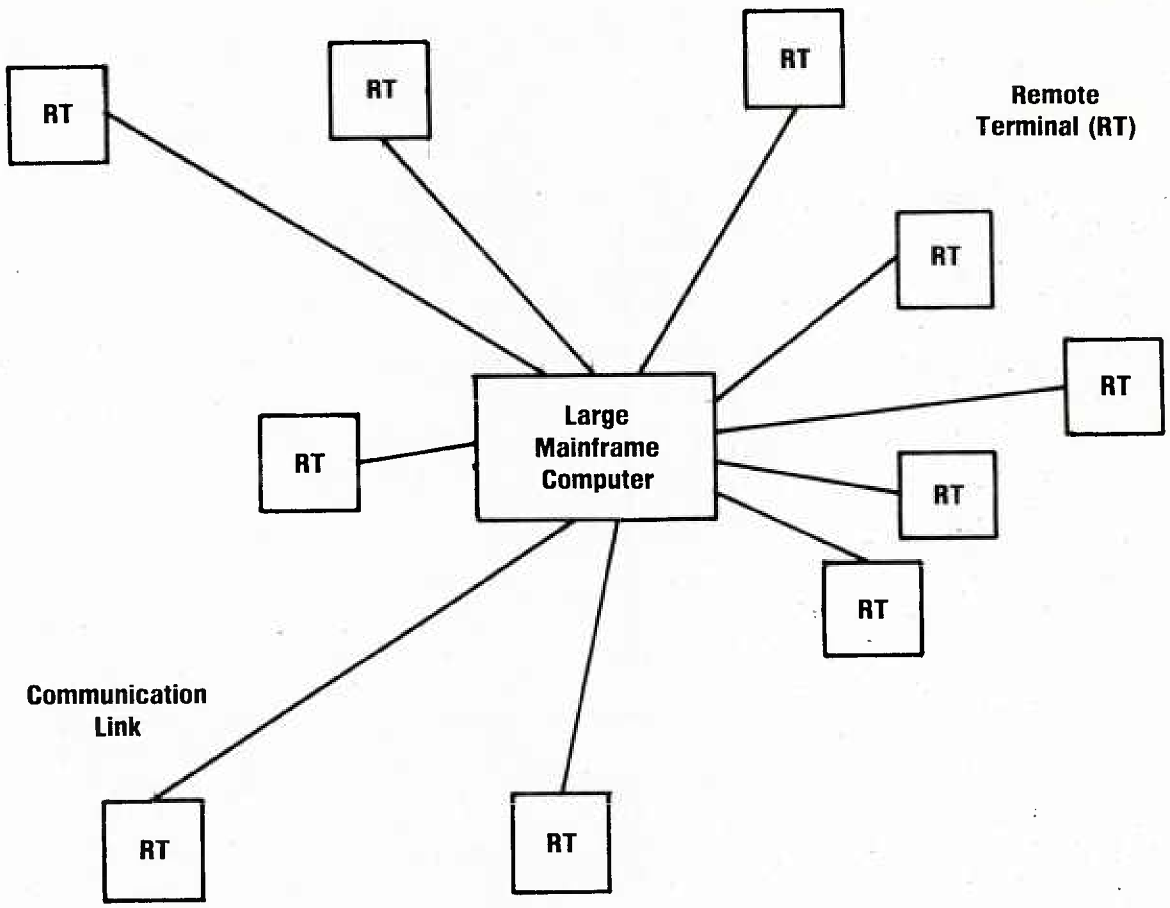


Figure 1. Centralized Large Computer Based System With Remote Terminals

PROPERTIES OF DISTRIBUTED SYSTEMS

Low cost minicomputers, microcomputers, and interconnecting communication links have made possible a new approach to information system design. This design approach is based on the concept of a distributed system which is a collection of small computers interconnected to allow communications and desired central control. The basic idea of a distributed processing system is to separate a large computing requirement into some combination of smaller computing tasks to be accomplished at the user's location. There are many distributed system configurations described in terms of networks. An entire section of this report is devoted to network architecture for some of the more widely accepted configurations. Two distributed system configurations are shown in figures 2 and 3.

In each of the figures there is a collection of minicomputers or microcomputers, a set of communication links, and an implied distribution of other resources and functions in order to enable the total system to perform its design mission. Note that the two configurations differ mainly in the way that communication is accomplished. This is an extremely important aspect of distributed systems and is dealt with in greater detail in section III.

Some of the advantages of distributed systems are:

- increased reliability and availability
- increased survivability
- distribution of processing power
- increased responsiveness
- increased modularity
- system expandability in smaller increments
- increased flexibility
- increased resource sharing.

The dispersion and distribution of computer processing power away from a central facility results in greater reliability and availability than would be obtained in comparable centralized systems. Because system components are relatively independent in a distributed system, a single unit failure will not cause a failure of the entire system. With a distributed design, the necessary computing power is always available to most of the users. For military applications, distributed systems offer the advantage of increased survivability. As long as communication links exist between several processing units in the distributed system, some level of service capability will always exist. The distribution of processing power is another significant advantage in that it provides faster and more flexible response to local organization needs.

Since there is some processing power at the user's location, many service requests can be satisfied locally. Other requests might require the use of processing power or stored data at another location and would, therefore, take take additional time. Modularity also impacts positively on system expansion capability, since small increments of capacity can be added quite easily and for relatively low cost. Another benefit is that an initial

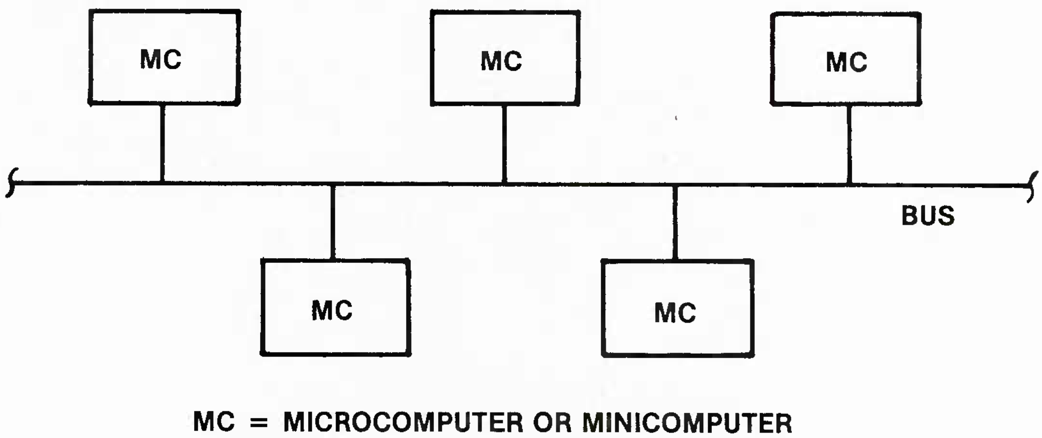


Figure 2. Hierarchical Distributed System

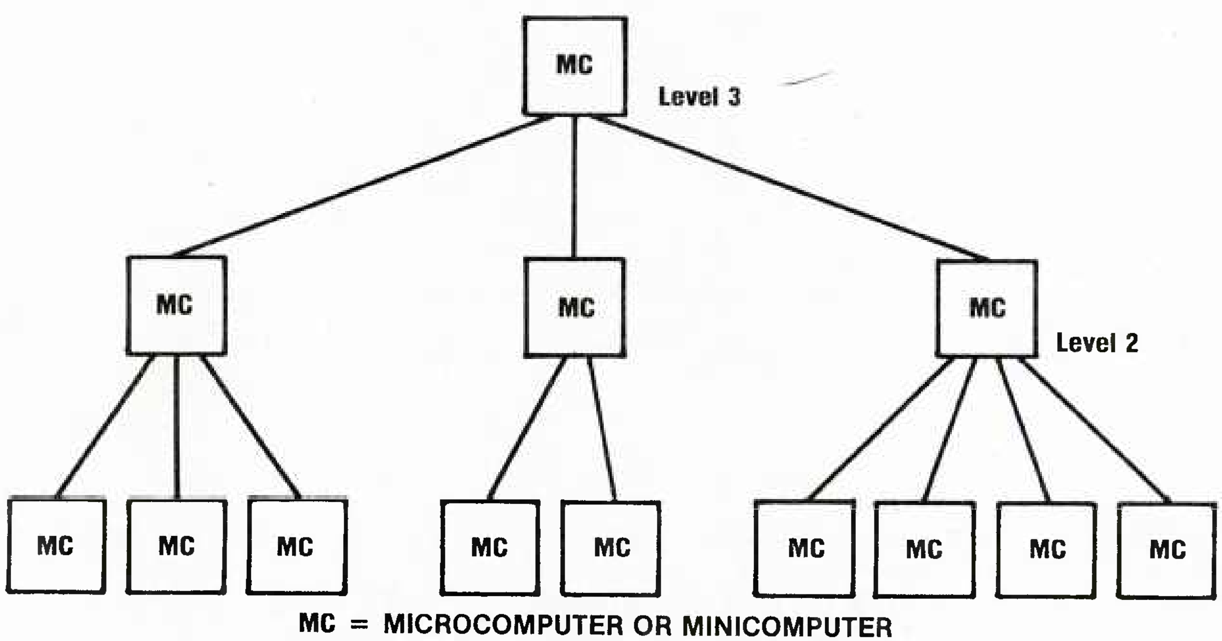


Figure 3. Multidrop Bus Distributed System

implementation of a distributed system can generally be achieved at a lower cost than a comparable centralized system.

System flexibility is also enhanced in a distributed system since the smaller components can be rearranged as necessary or added to when new or revised system functions are specified. At times, the dispersed processors can be performing highly individualized tasks, and at other times the entire collection of components can be working together to duplicate the capability of a large centralized system. Another factor that impacts on the cost of distributed systems involves the ability to share resources. It is possible to formulate designs which spread system resources. Large data storage facilities, fast printers, high resolution graphics terminals, and other equipment can be distributed throughout the system, while allowing any user to request the services of one or more of these devices. Use of shared disk memory is a common practice which allows a fast, large disk to be shared by many units, thereby resulting in lower total disk cost, compared to a system design using a large number of small, individually accessed disk units.

Distributed systems also have disadvantages, many of which prevent them from being the design of choice for all information processing systems. These disadvantages include:

- increased software complexity
- more difficult system test and failure diagnosis
- more dependence on communications technology
- unique expertise needed during design and development phase
- less system data security
- lack of total central control
- lack of standardization.

Distributed systems often require the production of software which is more complex than that required for centralized systems. The software is more costly to develop since each distributed processor may require its own operating system and must be capable of communicating with many of the other processors in the system. In addition, the software is likely to be more difficult and costly to test and debug since the distributed nature of the system causes testing and failure diagnosis to be more difficult. Although this situation applies to all system components, it is especially significant for software components.

Extensive dependence on communications technology, which is a necessary component of a distributed system, can also be a disadvantage. If the processors in the system are separated by great distances, the likelihood of a communications link failure increases. Also, if data transmission between system nodes is bursty, that is, if it occurs infrequently but is very intense when it does occur, then high speed communications links will be needed.

Such high speed links are more prone to problems than lower speed links and are much more costly. It is possible to realize decreased communication costs with distributed systems since, for the typical case, only summary data need be transmitted rather than extensive quantities of raw data.

The design and development of the hardware and software components of distributed systems may require the availability of unique expertise which at the present time is in limited supply. However, as more and more distributed systems are designed, developed, and implemented, more of this expertise will be available for new applications. A related problem in distributed system design and development is one of standardization or, more accurately, the lack of standardization. Mixtures of different types of processors, peripherals, data storage devices, and communications facilities may well occur in a large distributed processing system. To deal with this problem, system developers must have experience with a wide variety of hardware components and the expertise to integrate many of these components into a single system.

In a distributed system much of the system's activities will be accomplished at local sites in a relatively autonomous and independent manner. Under these circumstances, management control is a major concern. However, with good system design, this concern is minimized since the flow of management information to higher level nodes should not be impeded. Thus, top management will have the information needed to exercise control throughout the system. This control can also be extended to software configuration since applications software can be downloaded to run on remotely located processors. The problem of data security becomes more serious with distributed systems since sensitive information may exist at many remote locations. System security measures will have to be employed at all locations where sensitive data files are maintained.

NAVAL TRAINING DISTRIBUTED SYSTEM APPLICATIONS

The major focus of this report is on the use of distributed microcomputer systems for computer-based instruction with its major components of instructional/administrative support, instructional delivery, student management, testing, and curriculum development. The Navy currently uses several large centralized systems for some tasks related to computer-based instruction, but, in general, the use of such computer-based systems has been restricted because of the costs involved. The rapidly decreasing costs of microcomputers combined with the advances made in the use of this technology for training has led the Navy to an affirmative policy to use training automation whenever cost effective (CNETINST 3920.1B).

It may be instructive at this point to examine a typical training application to demonstrate some of the benefits of distributed systems. Consider a group of microcomputers used for individualized instruction student work stations. A single classroom might have 30 such work stations which are connected to a classroom management microcomputer with a high capacity disk memory. Several classrooms can, in turn, be connected to schoolhouse level computers with a much larger disk memory. It is then possible to interconnect one or more schoolhouses to higher levels of training management

with larger computers and very high capacity disk memory units.

The operation of the complete system is greatly decentralized, but a degree of centralized control exists to ensure the flow of management information on student and system performance. Training courseware produced at a development location can be sent by communication links to the school houses and then to the classrooms. This results in the storage of data files on classroom management station disks which can be accessed by microcomputers at the individual student work stations. Thus, individual stations do not require a dedicated disk storage unit, since they are able to share one large disk. When the basic courseware is downloaded to the classroom level, the operation of the classroom is virtually independent of the remainder of the system. As long as the classroom control microcomputer is functional, the individual student work stations can access the single large disk memory with the courseware for a given module or course. If the classroom control microcomputer should fail, the individual work stations can still function partially with the data stored in main memory, but they would not be able to obtain new data files until the control computer was available again. With identical microcomputers in the classroom, a physical exchange of the failed control computer with one of the individual work station computers could provide a means for keeping the system operational until the failed control computer is repaired.

Management information on the operation of the instructional system would flow from the lower levels to the top level, with aggregation and summarizing occurring before transmission. Thus, any intermediate level in the distributed system could get detailed performance data on all lower level activities. Class performance data on individual student progress would be monitored and recorded by the classroom control computers, and this data would be stored on the large disk memory connected to the control computer. At the schoolhouse level, the microcomputer could access the classroom disk units and get the complete file of data transmitted to its disk storage unit, or could get any part of the required data. Data files maintained by one classroom level disk unit could be shared among other classroom units by communicating through the schoolhouse computer. Similarly, the schoolhouse computers could communicate through the next higher level and share resources and data. Expensive peripherals such as graphics displays, high resolution plotters, and high speed printers could be dispersed among several classrooms and could be accessed as needed by any of the system processors.

At the schoolhouse level, the interconnection of a number of schoolhouses allows resource sharing and also allows the support of a common program development center. Management information at this level would allow evaluation of different classrooms for determining student progress, cost of instruction, and hardware system operational data for determining reliability and availability. The major focus of the Instructional Program Development level would be to produce courseware to use at the classroom level. The largest data base files would also exist at this top technical level. These files would allow for a complete synthesis of individual student progress, class performance, schoolhouse and school performance, and evaluation of the overall computer-based instruction system. The connection to the top administration level would be the last link in the hierarchy, and the equipment requirements

at the top level would probably be much less than that of the top technical level. Overall system management would be executed at this level, but control would be distributed through the system, making use of control capabilities at the lower levels. In essence, the control function would be maintained at a central level even though the operational functions would be carried out in a decentralized manner.

A distributed system can generally be tailored to user requirements in a much more flexible manner than a large centralized system. The expectation is that there will be many applications where a given distributed system design will offer more advantages than disadvantages, and will offer overall cost effectiveness for the design mission. The selection of a particular distributed system design requires evaluation of a whole range of performance trade-offs and cost factors. The remaining sections of the report are devoted to an in-depth discussion of the most important of these considerations.

SECTION III

NETWORK ARCHITECTURE

A distributed computer system can be described in terms of its network architecture. This architecture might be one of the generic forms, with a ring, tree, or bus topology, or it might be of an irregular mesh form with a unique topology. The network design is depicted diagrammatically as a collection of nodes which represent processors, peripheral units, or work stations interconnected via communications links. Various network configurations arise through different patterns of interconnecting the nodes of a distributed system.

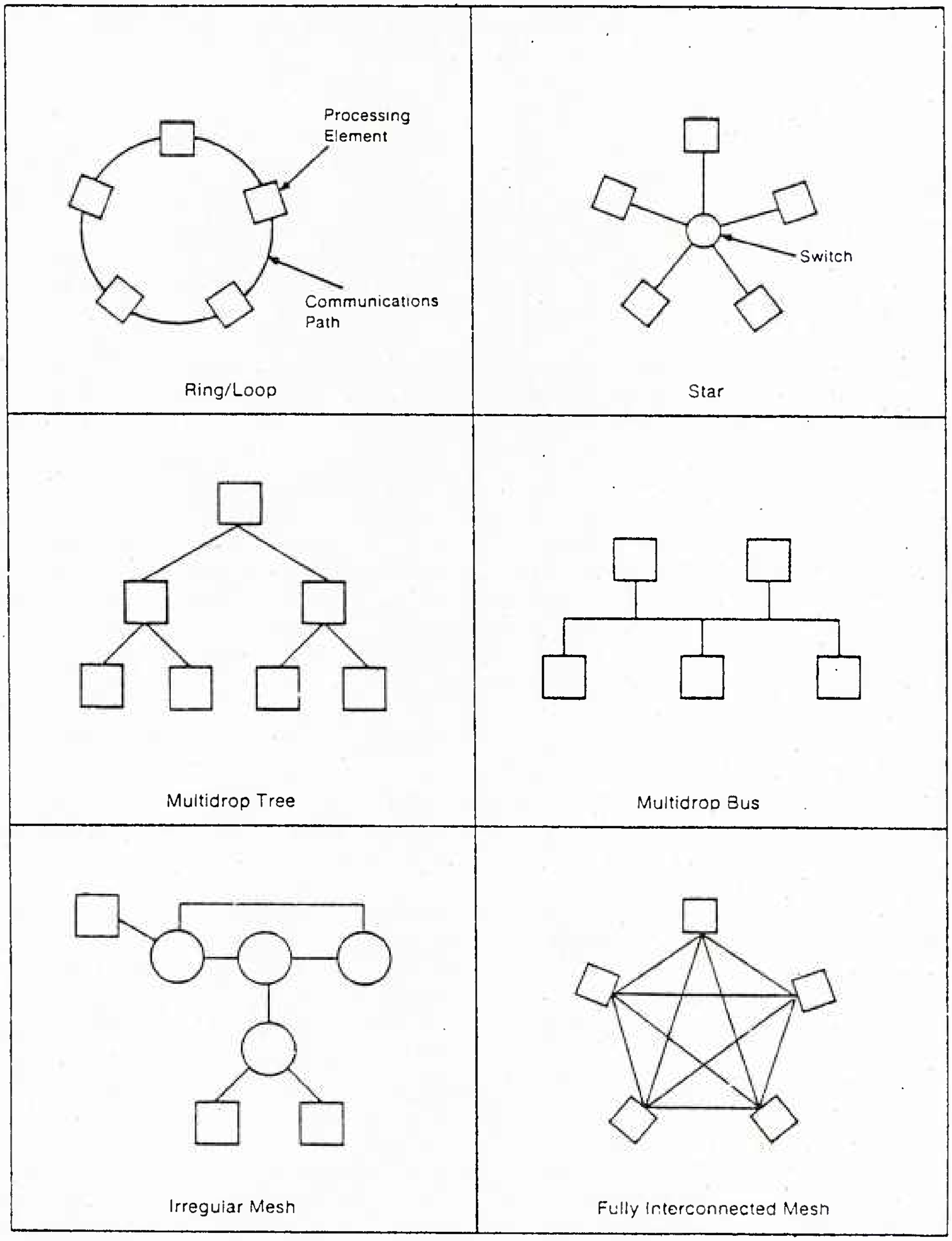
Many different classifications of computer network topology are possible but the generic designs shown in figure 4 provide an adequate framework for a discussion of the advantages and disadvantages of network configurations.

GENERIC CONFIGURATIONS

RING/LOOP CONFIGURATION. In a loop or ring network structure, a high-speed, one-way communication channel connects each node in an arrangement which forms a closed loop or a ring. The communication link may be a twisted-wire pair, a coaxial cable, or a fiber optics link. The loop or ring configuration is often used for a decentralized organizational structure in which coordination and communication between systems is required on an irregular basis. The loop or ring is usually limited to a small number of nodes or processors because average communication delay increases linearly with the number of nodes in the network. There are a number of advantages of ring/loop structures in terms of communications features and costs, and there is one major disadvantage -- reliability. Since a ring/loop is a serial structure, any single node failure stops communication on the entire network. A duplicate ring/loop can be used as a redundant component in the system to increase reliability but this becomes expensive and alters some of the advantages related to the simplicity of a normal ring/loop configuration.

Some of the advantages of a ring/loop network (Weitzman, 1980) include:

- ease of communication routing resulting from all nodes seeing all messages
- reduced need for modems since data transmission in the network is usually digital
- low network cost with total cost proportional to the number of node interfaces
- high communication rates because the nodal interfaces and not the processors are used to relay the messages, and more than one message can be in transit at one time



(© 1981 IEEE)

Reprinted with permission of IEEE (Freeman and Thurber, 1981).

Figure 4. Microcomputer Network Generic Topologies

- no central node overhead and delay since all nodes share equally in network communications.

The ring/loop network classification is based on the manner in which messages are communicated within the network. The three main varieties include the Newhall, Pierce, and Delay Insertion types. Further details of these different technologies can be found in Weitzman (1980). A relatively common term used in many ring/loop structures is "token passing," wherein a control message is passed from processor to processor. Only the processor currently in possession of the token can transmit a message, and any other processor desiring to send a message must wait until the token is passed around the ring/loop to its node. The operational detail of a ring/loop configuration can be quite complex. However, an in-depth discussion of network configuration design features will not be presented since they would go beyond the scope of this report.

STAR CONFIGURATION. In a star network structure one node forms the central facility for all communications between system nodes. This central node dependency offers some advantages and some disadvantages, since all communication switching is handled at the central node. Because of the expense involved, redundant central nodes are not generally practical. Additionally, if there is a great geographic separation between nodes, connecting cable costs can be very expensive since each node must be separately connected to the central node.

The star network is an example of a hierarchical configuration. Failure of a link between a given node and the central node will result only in the loss of service of that one node, but if the central node fails, all nodes relying on it lose service. A digital PBX system is an example of a large star network.

MULTIDROP TREE CONFIGURATION. In the multidrop tree or hierarchical network, the system has the shape of a pyramid with several levels of star configurations evident. The base of the pyramid, or the lowest level, is usually the location of the basic data interchange functions. In a computer-based instruction system, this level would consist of individual student work stations. The hierarchical structure is similar to a corporate or command organization wherein the lower levels perform well defined or specialized tasks and the upper levels have a more general purpose capability for controlling the entire system. This structure is readily adapted to the needs of a distributed computer-based instruction system where much of the operation occurs at the decentralized lower levels but where management information flow and central control is required at a higher level in the hierarchical framework.

The reliability of a hierarchical network can be affected in different ways depending on the location of a failed processor or communication link in the hierarchy. If a processor at the lowest level fails, or if a single communication link at the lowest level fails, then only one processor will be affected, eliminating the service contribution from that node of the

system. If that one processor were handling part of an overall task passed down from a higher level in the system, then that part of the effort would be lost. Thus, in the event of processor or communication failure at the lowest level, the total system capability is degraded but it is not necessarily lost. It is likely that only a single user would experience loss of system service. As failures occur at higher levels in the hierarchy, more of the system is adversely affected. If the system has four levels, and a level three (next to the lowest) processor fails, then all of the level four nodes connected to this processor would lose access to higher level nodes. If the lowest level nodes are performing decentralized tasks, they can continue doing so as long as data or instructions are not required from the higher levels. In general, the nearer the top of the hierarchy a failure occurs, the greater the number of nodes in the system that are potentially affected.

MULTIDROP BUS CONFIGURATION. In the multidrop bus network structure the connected processors communicate over a common channel and a message flows from the originating node in both directions to the end of the bus. Physically, a multidrop bus can be viewed as a set of nodes connected to a coaxial cable which is terminated at each end by a special matching device. As a message flows along the bus, each node reads the message address and determines if the message is to be received. The bus configuration has the simplicity properties of the ring/loop network, but it does not suffer from the same type of reliability problems. With a global bus network, the bus control logic is distributed throughout the processors connected to the bus, and there is no central control facility which is critical to the entire operation of the system. Bus architecture is widely used in aerospace, industrial, and laboratory automation applications (Weitzman, 1980). For many applications the bus architecture allows the simplest, lowest cost, and most reliable system design.

There are many bus structures for centrally controlled and decentrally controlled networks. For central control, there is a polled bus, an interrupt driven bus, and a slotted bus structure. For decentralized control, there is a global frequency division multiplexed bus, a global time division multiplexed bus, and a global multiple access bus. A detailed description of each of these particular bus structures can be found in Weitzman (1980). A well known example of the global multiple access bus is the Ethernet developed by the Xerox Corporation. The Ethernet consists of some 100 nodes connected to a one-kilometer long coaxial cable. In general, the physical material for a bus could be a twisted pair of wires, a multiwire cable, a coaxial cable, or a fiber optics cable. The speed of transmission of data along a bus is greatly affected by the transmission medium used.

The reliability of a bus network is dependent on whether the bus is under central or distributed control and on the number of shared units connected to the bus. If a given bus is very long, it might require the use of repeaters to amplify the signals on the bus; the existence of such repeaters lowers the overall system reliability. The form of the transmission medium also affects the bus network reliability, with a coaxial cable being more reliable than unprotected twisted pair wires. In general, the

reliability of a given bus structure will be better with the simplest approach to control logic, and the minimum use of repeaters.

IRREGULAR MESH CONFIGURATION. The irregular mesh network structure is a general, unconstrained topology with nodes connected together in an arbitrary manner. Not every node will be connected to every other node, so this network is not fully interconnected. This network is a partially connected subset of a fully connected network where a sufficient number of interconnections is available to make use of a communication technology for circuit, message, or packet switching. The irregular mesh configuration is used in many long-haul packet communication networks and is typically applied to efficiently utilize expensive transmission media. However, long message delays, which can frequently occur, cause system response time problems. Little general use of this network topology appears to exist. However, one well known application is the ARPANET system developed by the Advanced Research Projects Agency of the Department of Defense. The most prevalent application of this partially interconnected mesh network, also called a point-to-point network, is to connect geographically dispersed computers through a common-carrier communication circuit. Parts of a distributed computer-based instruction system could use similar design concepts if there are substantial distances between major nodes of the network.

FULLY INTERCONNECTED CONFIGURATION. Fully interconnected networks have a dedicated link between each pair of nodes in the system. Such a system design is impractical for a large number of nodes. As a practical matter, fully interconnected networks are rarely used when more than five nodes exist in a system. One advantage of such a network design is that software is not required for making message routing decisions since all nodes are always connected to all other nodes. Another advantage is that failure of a single communication link has a very limited affect on overall system performance. However, as previously discussed, the cost of this network configuration is very high for a large number of nodes, since N(N-1)/2 communication links are required for a system with N nodes.

The use of the fully interconnected network topology is generally restricted to special applications. One university application of this network design connects a variety of small processors and peripherals to form an interconnected system with extremely high data processing and input/output data rates (Freeman and Thurber, 1981). The software costs in this application were minimized since separate communication routing programs did not have to be written for each separate computer.

GENERAL RECOMMENDATIONS

Even within the category of computer-based instructional systems there is such a diversity of potential applications that no single set of guidelines for preferred distributed processing networks can be provided. However, general guidelines can be given to aid in the conceptualization of alternative system designs to satisfy specific applications.

At a classroom or schoolhouse level a "local network" would be a preferred design to consider. A local network is simply a computer network which is restricted to a geographic range of a few kilometers or less. An entire section of this report is devoted to local networks because of their expected impact on computer-based instructional system designs in the future. Local networks almost exclusively use the ring/loop or the multidrop bus configuration since these are simple networks with reasonably low cost and relatively high reliability. A star configuration is also viable as an option, but there are currently only a few commercially available local networks using this design.

The hierarchical network configuration will best serve the needs for most computer-based instructional systems. In a typical computer-based instructional system, there are individual user work stations at the lowest level of the network, and there is a need for some level of central control at higher levels of the network. In addition, there is usually a desire for a flow of management information from lower to higher levels. The overall operation of a computer-based instructional system is decentralized in many respects, yet it requires some level of central control less than total centralization. The hierarchical network structure matches this system operational procedure very well and it is not inconsistent with the use of a local network at the lowest level.

If nodes of a computer-based instruction system, such as schoolhouse nodes, are geographically separated, then a partially interconnected mesh configuration might be most appropriate to make use of common carrier communication lines. Limited applications of this network structure or the fully interconnected mesh structure are envisioned for use with computer-based instruction systems. It is also possible that applications at the classroom or schoolhouse level might benefit from use of the fully interconnected mesh network in cases where a group of processors are connected to a large number of peripheral components.

The design of a distributed system network architecture is heavily influenced by communication requirements, capabilities and costs as well as by the distribution of resources over the network. These topics are covered in the next section of this report. A separate section is devoted to the communication aspects of local computer networks.

SECTION IV

RESOURCE DISTRIBUTION

Computer resources typically consist of processors, data storage units, input/output controller devices and communications equipment. In this section the distribution of these resources within a distributed processing network will be addressed. It is useful to consider system resources as being distributed over levels or layers of the system rather than just in a physical or geographic sense. The layers or levels are related to the functional requirements of the system design and also determine the interconnections of the various resources. This concept applies to both hardware and software and serves to facilitate good design modularity.

The term "system resources" is a broad label which is meant to allow a wide range of hardware devices and software capabilities. Some examples are microcomputers, minicomputers, CRT terminals, disk drive units, communications processors, high speed printers, and tape drives. The ability to substitute one network component for another and still satisfy the system design requirement is provided by modularity -- the dispersal of resources among the various nodes of the network. The three major concerns in resource distribution involve the (1) amount of processing power to place at each node in the network, (2) portion of the total system data base to be located at each node, and (3) communication capacity needed at each interconnected node-pair.

PROCESSOR DISTRIBUTION

Recent distributed system designs incorporate microcomputer and minicomputer components that have large processing capacities and low purchase costs. A typical system will also include terminal devices which are system extensions to these new computers. With the emergence of 16-bit microcomputers and 32-bit minicomputers there has been a real blurring of the original terms micro and mini. Many present day minicomputers have equal or greater capabilities than a mainframe, general purpose computer of several years ago. New 16-bit microcomputers have most of the capability of many of the general purpose computers used 5 to 10 years ago. Thus, it is very difficult to talk about processing power needs in terms of the generic categories micro, mini, and mainframe computers.

There are, however, significant differences in capabilities and costs of these computers, and future specifications may address these differences in quantitative terms rather than in the general qualitative terms used today such as mini, micro and mainframe. Capabilities of the newer 16-bit microcomputers, such as the Intel 8086, Zilog Z8000, Motorola 68000, and the National NS 16000 family may include clock rates of 10MHz with up to 64 megabytes of addressable memory. These are capabilities which most people would classify as mainframe characteristics. Regardless of the actual capabilities of a given central processing unit (CPU), the main determinant of its cost and performance is the additional hardware connected to it, such as memory and input/output ports. A $10 microcomputer chip may have a tremendous potential capability, but it may only utilize a small portion of this capability if it is not augmented with additional hardware. Thus, a user terminal

may contain a powerful microcomputer chip but still be relatively "dumb," or it could act as a total "stand-alone" computer if properly configured with additional components.

The development of specifications for the processing power needed at each node of a distributed system network can be aided by examining three general categories of functions performed by an information system. These are (1) data collection, (2) data processing, and (3) data base management. The data collection task is usually the largest function of an information system, and, in a hierarchical system structure, this occurs at the lowest level where the users are. In a computer-based instruction system, this activity would take place primarily at individual student instruction stations. The data processing task may involve almost any form of computing or problem solving, file manipulation, or other activity to rearrange data in the system. Scoring examinations and prescribing remediation work are examples of data processing in a CBI system. Data base management involves the creation and maintenance of an organized set of files for information storage and retrieval. In a CBI system, these files would typically include course modules, student records and test item data. Given these three task functions, and the system operational and design goals, it should be possible to propose several alternative system configurations to meet the overall objectives.

There are some other items of note within the category of processor distribution which relate to systems of homogeneous or heterogeneous processors; i.e., are all the processors identical or is a mixture of models and types allowed? A homogeneous system has an obvious advantage of common processor hardware which means only one type of computer needs to be maintained and serviced. Cost savings occur because of similar spare parts and reduced maintenance training requirements, and system reliability is enhanced because one failed computer can be physically replaced with any other computer in the system. Software costs are also minimized with one common computer and programmers need only be familiar with one piece of hardware. A common computer programming language can also be used throughout the system with complete compatibility. The disadvantages which can accrue result from the rigidity of system configurations to maintain homogeneity.

The general approach to specifying the distribution of processing power in a distributed system network is one of selecting a system which has the potential for satisfying the requirement and then trying it in the operational setting or in a simulated setting. This approach is normally influenced by state-of-the-art hardware availability and capability. As hardware capabilities continue to increase and costs continue to decrease, the choices become easier. The approach can be to put a highly sophisticated processor at each node and add only the needed amount of supporting hardware to meet the basic requirement. This design methodology may turn out to be the most cost effective when contrasted to more formalized design and development procedures.

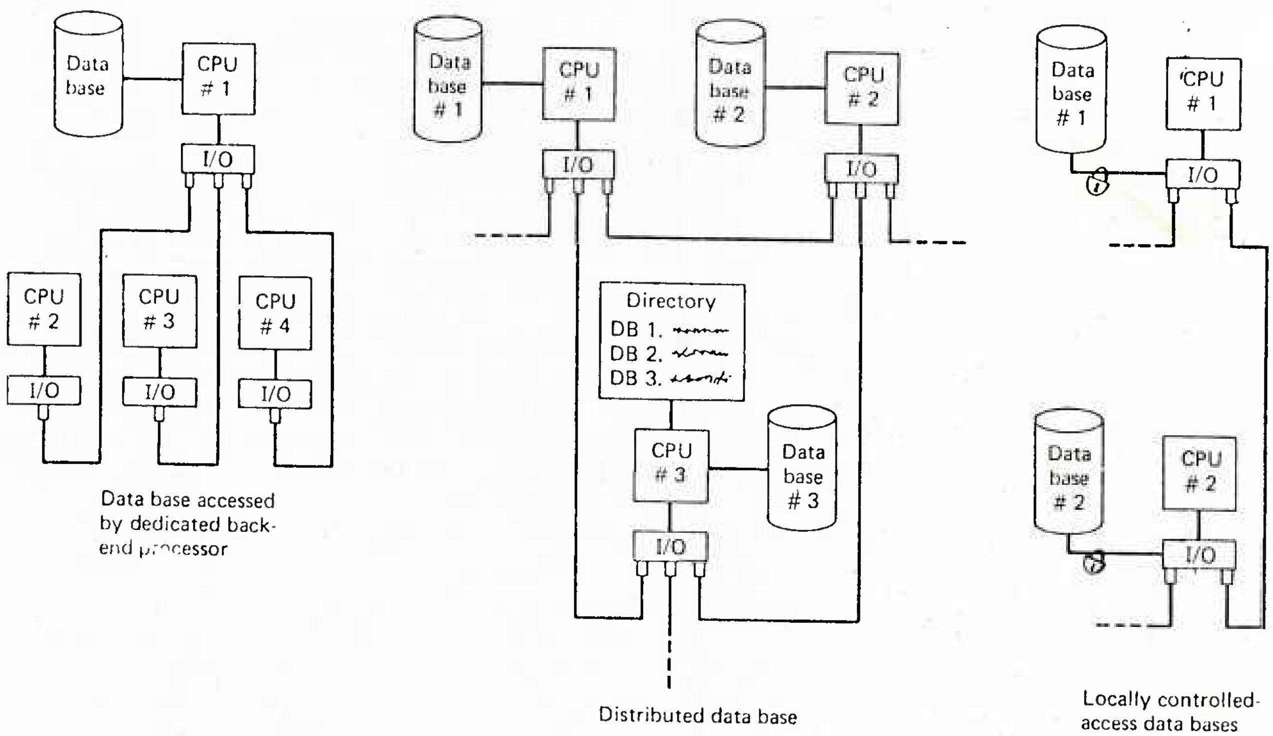
DATA BASE DISTRIBUTION

A data base is a collection of logically organized data stored for use in one or more applications and is intended to be accessed by multiple users. A single user would not normally require the service of the entire data base but only a portion of it that is specific to a given task. In a CBI system the data base would contain data for course organization, instructional program delivery, administrative recordkeeping, and report generation. The method used in distributing the total data base over the many nodes and processors in the system is a major consideration in distributed system design. There are three possibilities for locating data files in the system: (1) centralized, (2) local or fully distributed, and (3) part centralized and part local. Some of these approaches are shown in figure 5.

In a centralized data base configuration, a complete set of files resides in a central facility; and this data base arrangement is independent of the system configuration which can be centralized or distributed. Some of the benefits of a centralized data base are elimination of data duplication, ease of file standardization, improved data security and control, and single staff data base management. Arranging the data base in a decentralized manner requires the partitioning of files into pieces that can be accessed by a number of dispersed processors and users. The advantages of distributing the data base include better system reliability and availability, improved system response time because of file collocation with users, improved data security and privacy, and reduced communication requirements if system users are geographically separated.

Another important factor in the consideration of data base distribution is whether files will be partitioned into unique pieces or whether they will be redundant or replicated. In the case of a replicated data base, all or part of the data base is copied and stored at two or more nodes of the distributed system. With a partitioned data base, there is no redundancy or replication, and the total data base is the sum of all of the independent parts. In a CBI system the student files could be partitioned so that they only exist at individual schools or they could be replicated and stored at the school and the school command location.

A replicated data base takes more storage capacity than a partitioned data base, but there are some advantages to be gained by incurring this extra cost of storage. One advantage is greater system reliability and availability. If a failure occurs in the storage and retrieval function for a given file, then a copy of that file which is replicated somewhere else in the system can be accessed. The alternative is to wait for the unique data file to be available, and possibly to reload it from some master backup copy and update it if some catastrophic failure occurs. Some reduction in communication cost may also occur if part of the data base is shared by multiple users.



Reprinted with permission of Prentice - Hall (Weitzman, 1980).

Figure 5. Approaches to Data Base Handling in Distributed Micro- and Mini-Computer Systems

Another factor closely related to the data files in a system is that of the directory -- the listing which shows where a physical file or record is stored. In a centralized system, the data base and the directory both reside in the same computer facility. For a distributed system, the location of the directory is somewhat analogous to the location of files. Typically, it would be either replicated or divided into unique pieces. Use of a distributed directory involves copying the entire directory for the complete data base at each node in the system, as opposed to a local directory where only that portion of the directory related to the local data base is stored at a given node. The location of the directory can affect system availability, response times, communications volume, and overall system complexity.

The functional requirements for a distributed system can be used in determining the best data base location. A system which has a widespread geographic form may be such that the data base can be partitioned or replicated in a manner similar to the geographic dispersion. If separate functions are involved at the different geographic locations then a partitioned data base responding to these separate functions might be appropriate. A hierarchical system structure is amenable to supporting a hierarchically partitioned data base. In a hierarchically structured CBI system, for instance, the basic student performance records can be located at the classroom level without these files being replicated at higher levels in the system. Information is communicated upward and downward in this CBI system, and upper levels must collect data from lower levels to prepare comprehensive reports. Another data base structure for a CBI system could use hierarchically replicated files, in which classroom level student record files are duplicated in a master data base at a higher level. The most general data base structure would be a combination of partitioned and replicated data. In this case, certain files would be partitioned among the nodes of the system, while other files would be replicates or partial replicates of certain of the partitioned files. In a CBI system, this might mean that partitioned student records are located at the classroom level and that course module data could be contained at both the classroom and the school level, being at least partially replicated. The use of trial system designs employing alternative data base structures is part of an overall design process for distributed systems.

COMMUNICATION DISTRIBUTION

Communications technology for distributed systems will be discussed in depth in section V. However, it is useful at this point to discuss some system design considerations relating to communications. The overall goal for the communications design is to identify communication methods and performance requirements which allow the best utilization of the entire distributed computer system, while minimizing communication costs. This requires a knowledge of the system performance in terms of speed and capacity requirements for all communication tasks in the system. Management information and control requirements must be clearly specified in order to determine the types of communication needed and the capabilities of the various communications channels. Data security is also a factor which must be examined at each level of the distributed system.

Distributed system resources are linked together by communication facilities which can vary greatly in speed, capacity, and cost. Components that are located in physical proximity may be connected through a local network. Geographically separated facilities may be linked through telephone lines, microwave links, or satellites. The communication processing hardware and associated equipment needed for a distributed system includes (1) transmission facilities, (2) coupling facilities, (3) multiplexors, and (4) communication processors. The transmission facilities include a wide range of technologies using electrical or optical modes such as telephone lines, fiber optic cables, coaxial cables, microwave links, and satellites. The coupling facilities are specified by the transmission mode, and interface the processors to the transmission facilities. An acoustic modem is a typical example of a coupling device. Use of multiplexors is almost certain since they allow a collection of low speed communication links to be integrated into one single high speed link. The communication processors perform the decisionmaking function in the overall communication network, and they handle switching and data concentration as well as providing an interface between the information processors and the other resources of the system.

Two of the major goals of the communication structure are to provide the necessary data paths between nodes of the system and to support the system response time and throughput goals. The total system design goals must be clear so that the communication capability can be distributed in such a manner to accomplish these goals. In a CBI system, it is very important to know the requirements for management information that must flow up and down the system in order to specify the communication structure. For example, the response time for passing a course module data file from a school command level to a schoolhouse level is directly affected by the speed and capacity of the communication link between these two levels. With a low speed line the response time could be an hour, whereas a high speed line would cut this time to minutes. The selection of the communications structure is an important part of the overall process of distributed system design and care should be taken to assure an effective communications system configuration to satisfy system needs.

RESOURCE DISTRIBUTION DESIGN CONSIDERATIONS

The selection of the processor, data base, and communications components for a given system is best accomplished by a process of total system design. This task is immensely complex because of the interactive nature of the components within the distributed system. For a given partitioning arrangement, the system data base will affect processor requirements and communication requirements and these in turn affect each other. There is very little chance of effectively accomplishing a "one-shot" design in terms of a logical progression of steps from some basic assumptions. The actual design process will most likely be a repetitive or iterative procedure where some basic assumptions are made and a trial design is formulated. The trial design is then analyzed and in most cases will require additional design effort. The system design is then altered and a new evaluation is performed. This trial and error process, although appearing unscientific, is the most systematic approach available at this time. Booth (1981) has formalized this procedure

as an algorithm for total system design. The first part of this procedure involves formulating a trial system design using the following steps:

1. select an information-processing structure or network architecture

2. select a database structure

3. select the types of components to be used for processing and database purposes

4. allocate information-processing functions to components

5. allocate database elements

6. select a structure and components for the interconnecting network.

Once this trial design has been formulated, it must be examined to determine its properties and its desirability. An iterative procedure consisting of the following four steps is then employed:

1. formulate a trial design

2. analyze that trial design in order to determine whether functionality, performance, and availability characteristics are adequate and whether costs are acceptable

3. if all parameters are not satisfactory, modify the trial design and repeat step 2

4. if a given trial design cannot be modified satisfactorily, return to step 1 and try a different trial design.

Booth also states criteria for evaluating trade-offs in alternative system designs. She recommends asking the following questions before deciding on a particular alternative:

1. Does the trial design provide the required functions?

2. Will it handle the required volume, with adequate capabilities for expansion?

3. Will it meet the response-time requirements of users?

4. Will it provide the minimum required level of availability?

5. Can the cost of the system be justified (or is it within prestated limits)?

Other general considerations in distributed systems design involve the availability of equipment that is experiencing rapid technological evolution

and attendant price reductions. Constant monitoring of new technology is critical during development of a new distributed computer system if it is to offer the necessary functions at the lowest cost. Technology forecasting may be helpful during the initial phases of a large scale distributed system design although it is likely that any design, when converted to operational equipment, will be behind the technological state-of-the-art.

FUTURE PROJECTIONS

Improvements in hardware capability and reductions in hardware costs will most likely lead to distributed system designs wherein an intelligent processor such as a microcomputer and a portion of the system data base stored on a hard disk will reside at each node in the system. In a CBI system with individual student work stations there may not be a need for a hard disk at the work station since the RAM memory may be adequate. However, at all other nodes there will most likely be a need for both the processor and the disk. The exact nature of the capabilities of each processor and disk cannot be specified, but it is likely that the new 16-bit microcomputers available will be able to provide most of the processing power needed. A CBI system design will be most appropriately served by a hierarchical network structure, with the communications specification based on management information flow and student station information flow requirements. The individual student stations will probably be connected to a control processor through a local network which has the speed and capability to provide the required response times for all instructional functions. The remainder of the communications links will have speeds and capacities dictated by higher level management control function requirements.

SECTION V

COMMUNICATIONS

This section provides a general treatment of the terms, technologies, and operational considerations of communications as they apply to use in distributed computer systems such as those for CBI. Without the communications links, the entire system would be simply a collection of individual pieces limited to performing those tasks within the processing capabilities existing at single nodes. The development of a total distributed system capability requires a high degree of sophistication in the design of the data transmission facility.

The major components of a computer communication system include transmission facilities, coupling facilities, multiplexors, and communication processors. The transmission facilities may be telephone lines, fiber optic cables or coaxial cables, and the coupling facilities may be modems or other coupling devices which interface communication or information processors to the transmission lines. Multiplexors are devices for combining a number of low speed data links into one high speed data link; communication processors are computers for concentrating data, switching data paths, or interfacing information processors to the distributed network.

TRANSMISSION MEDIUMS

A wide variety of transmission mediums exist, and it is not useful to give an exhaustive list here. However, there are two general categories of interest: electrical and optical. Electrical mediums are either in the wire or radio class; with wire-based technologies being twisted-pair wires, general cables and coaxial cables; and radio-based technologies being broadcast facilities or point-to-point facilities such as ground microwave links and satellite microwave links. A special class of wire technology of great interest is dial-up (switched) or private, leased telephone lines. Many distributed systems rely completely on the use of telephone lines for their data links.

The use of a particular communication medium is dictated by factors of cost, availability, physical durability, speed of transmission, nature of the information (i.e., analog, digital, video), and data security requirements. It is likely that a mixture of mediums will be used in different parts of a single distributed system. For example, in a CBI system with a hierarchical configuration, the lowest level nodes may be connected through a local network using a coaxial cable, while the next level up may be linked with a dedicated telephone line, and the next level up may be linked through a communication satellite.

Optical-based technologies are less frequently used than electrical-based technologies in general, but developments of optical fibers are producing interesting capabilities at costs that are declining rapidly. Optical fibers offer some unique advantages in that they are small, can support very high data rates, and are not affected by electrical and electromagnetic noise or interference. Fiber optic cables are most

applicable to short computer-to-computer connections or for local computer networks. This technology will probably not be cost effective until 1984-86, but the time frame is clearly of interest for development of future distributed computer systems. Laser technology is the other major optically-based medium, and it is also a technology of the future, with anticipated cost-effective applications in the late 1980s.

One of the major factors to be considered in selecting a transmission medium is the speed or the data rate of transmission. The standard unit of rate is the "baud," which in many instances is simply "bits per second." To understand what this means we must select a data encoding method to use as an example. The ASCII code is a common scheme for data transfer between computer facilities and uses seven data bits to describe a character such as a letter, number, or control character. Using synchronous transmission (described shortly), 10 bits are used: seven for the character code for ASCII, plus one start bit, one stop bit, and one parity (or error checking) bit. Thus, a 300 baud rate, which is a common voice phone line rate, translates into a data rate of 30 characters per second. To put this into perspective, a standard page of 250 words of 5 letters each, which would give 1,250 characters, would require just over 41 seconds to transmit at a 300 baud rate.

Data rates for a variety of transmission mediums are shown in table 1.

TABLE 1. DATA RATES FOR COMMON TRANSMISSION MEDIUMS

Transmission Medium	Data Rate
Dial-up telephone line	300 baud asynchronous 1200 baud synchronous
Leased telephone line (conditioned)	2400 baud 4800 baud 9600 baud
Twisted-pair wire	250-1000 K bits/sec
Coaxial cable	
Baseband	4-10 M bits/sec
Broadband	10-250 M bits/sec
CATV Broadband	10-400 M bits/sec
Microwave link (Video channel)	6 M bits/sec
Satellite link (Video channel)	6 M bits/sec
Fiber optic cable	50-? M bits/sec

DATA TRANSMISSION AND COUPLING

Computer data are usually sent over a transmission line in serial form, which means bit-by-bit, sequentially. A pictorial representation of several computer communications configurations is shown in figure 6. The modes of transmission may be asynchronous or synchronous. With asynchronous transmission, each character is bracketed with a start code and a stop code, and then sent as a complete unit. Other bits, such as a parity or check bit, may be added to the basic character before it is enclosed with the start and stop codes. With synchronous transmission, a block or group of characters are transmitted without start codes or stop codes, but several synchronizing bits are sent periodically to keep the transmitter and receiver on the same timing basis. If data occurs only intermittently, then asynchronous transmission may be preferable since characters can be sent as they occur. Since blocks of characters are sent with synchronous transmission, individual characters must be placed in a buffer, or temporary storage, until enough are accumulated to send them. Thus, synchronous transmission equipment requirements and cost are greater than those of asynchronous, and are only justified if large quantities of fairly continuous data occur.

Because of practical operational aspects of the telephone system, it is not possible to send digital signals directly over a telephone line--even a conditioned, leased line. The digital signals must be converted to audio, or analog, signals before transmission over a voice telephone line. This conversion is accomplished by a modulator at the transmission end and a demodulator at the receiving end of the line. Since data are usually sent and received by a given facility, there is a need to have both the modulator and demodulator at the same location. A combined device which performs both these functions is called a "modem," where the term comes from the MO of modulate and the DEM of demodulate. Modems are designed for a given data rate, which accounts for the common modem rates of 300, 1200, 2400, 4800, and 9600 baud. A given modem will operate either asynchronously or synchronously, with a typical low speed, low cost modem providing 300 baud asynchronous transmission over regular dial-up telephone lines.

Another aspect of data transmission is the direction of transmission in terms of sending, receiving or performing both simultaneously. A simplex line, or mode, allows transmission in only one direction at a time. A half-duplex line, or mode, allows bi-directional transmission, but not simultaneously. The full duplex line, or mode, is the most flexible (and costly) and allows simultaneous transmission in both directions. Both equipment and line costs usually increase as the capability for half-duplex or full-duplex is achieved. System design requirements need to be carefully examined to ensure that the necessary capabilities are there but also making sure that excess capability is not paid for when it is not needed.

Multiplexing and concentration are two other data transmission functions that are performed in distributed computer systems. A multiplexor combines two or more low speed data streams into a single, high speed data stream, and does so using either time division multiplexing (TDM) or frequency division multiplexing (FDM). TDM is more efficient and more costly than FDM.

Figure 6A General Communications Model

Figure 6B Digital Communication Between Processors

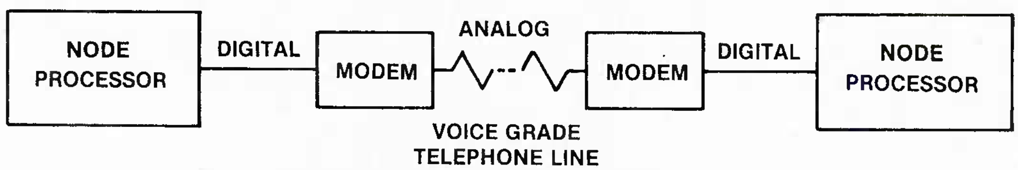


Figure 6C Need For Modems

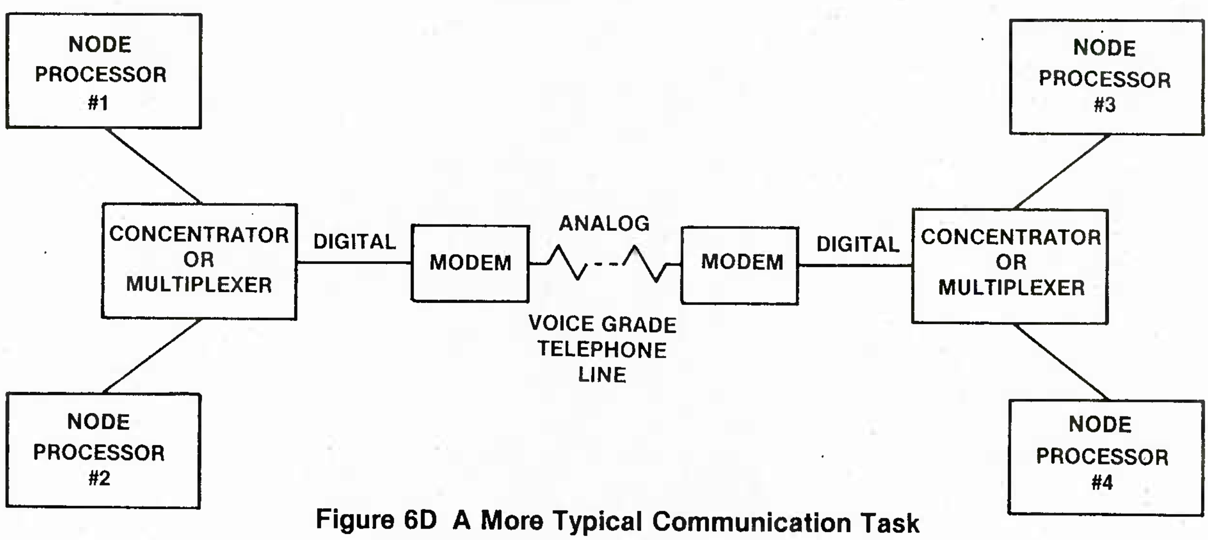


Figure 6D A More Typical Communication Task

Figure 6. Computer Communications Configurations

FDM is simpler and generally used for modems with a data rate of 300 or 1200 baud. Concentration is a process similar to, but somewhat broader in purpose than, multiplexing. A concentrator is usually a computer or processor which stores data, schedules data transmission, and allocates a larger number of slow speed data streams to a smaller number of high speed data streams. So, a concentrator performs the function of a multiplexor, but also provides an intelligence and decision-making function to produce fast and efficient transmission of data.

COMMUNICATION PROCESSING

In a large computer communications system there are a great number of decisionmaking functions related to communications tasks, which must be performed by an intelligent device. Commonly, a communication processor (or network processor, or front-end processor) accomplishes these functions which typically include multiplexing and concentration tasks. The general functions performed by such a communication processor include modem control, multiplexing, concentration, error detection and recovery, code conversion, network interfacing, protocol handling, and message routing and switching. Every computer network must have a set of rules and procedures for sending and receiving messages between nodes, performing system functions, and managing data flow throughout the system. These functions may be accomplished with a dedicated microcomputer or included as part of the overall function of the processor at a given node in a distributed system.

Two major areas of interest in communications processing are the message protocols used and the message routing and switching functions.

NETWORK SWITCHING

A distributed computer system does not typically have a collection of totally dedicated communication facilities linking all nodes. Therefore, some form of network switching is required. The two major approaches to network switching are direct data transmission and concentrated data transmission. With direct data transmission, circuit switching is used to establish a physical electrical connection between the data source and the receiver, and that connection is maintained until the transmission of data is completed. A standard voice grade, dial-up telephone line is an example of a circuit switched network. With concentrated data transmission, often called store-and-forward communication, a message is placed in a buffer for temporary storage prior to actual transmission. The two most common approaches to this technology are message switching and packet switching. With message switching, the entire message is accumulated in a temporary memory and then transmitted between nodes in the network when the communication link becomes available. There may be temporary storage at several nodes in the network until the message finally reaches its destination. With packet switching, a message is broken up into fixed-length pieces called "packets," and these pieces are individually transmitted over the network. A typical packet size is 1,024 bits, or 128 characters. Each packet is routed over the network to its final destination, and different packets may take different routes.

Since it is possible for packets to arrive at the final destination out of sequence, the node processor must be able to reassemble the complete message in the proper packet order.

The Department of Defense ARPANET was the first large-scale packet switching network. In 1979, over 100 computers of many varieties were connected throughout the nation. Packet switching networks generally provide redundancy through multiple path connections and thus have a high degree of availability. The CCIT X.25 protocol (discussed in detail later in this section) is available as a standard for connecting systems into these packet switched networks. The use of packet switching is often justifiable when data flow is intermittent, and data volume is large. Some comparative features of the various types of network switching are shown in table 2.

TABLE 2. PERFORMANCE CHARACTERISTICS OF COMMON NETWORK SWITCHING TECHNOLOGIES

Communications Feature	Circuit Switched Network	Message Switched Network	Packet Switched Network
Transmission Setup Delay	High	High	Low
Network Switching Delay	Very Low	High	Low
Transmission Link Utilization	Low	High	High
Reliability	Low	Low	High

Source: McGlynn, 1978.

Data security is a great concern in terms of transmitting large amounts of sensitive information over long distances and through networks accessible to a large number of people. This is a topic of such broad interest and concern that it will be addressed in a separate part of this section.

PROTOCOLS

A communication protocol is essentially a set of rules that controls the communication system operation. These rules are necessary to ensure compatibility between sending and receiving parts of a computer system which must communicate information such as, "you can send now," "I received the information," and other similar control and format information. The term "protocol" is broadly defined by the information processing community to include physical and electrical compatibility criteria as well as criteria

for the actual transfer of information. According to the protocol definition adopted by the International Standards Organization (ISO) there are seven levels or layers which specify these different criteria. Lower level protocols, such as the RS-232 interface standard, specify the physical and electrical compatibility requirements while the higher level protocols facilitate the flow of messages between the various parts of a computer-communications system. The Department of Defense ARPA Network, or ARPANET, for example, defines four layers of protocols.

The general levels or layers of a protocol include those for (1) physical interface, (2) electrical interface, (3) communication link control, and (4) message handling. These first two levels are usually specified by industry standards such as the Electrical Industries Association (EIA) RS-232C for serial communication. More recent EIA standards for RS-422 and RS-423 are expected to replace the RS-232C standard eventually (McGlynn, 1978). The higher protocol levels for communication link control and message handling must specify the formats for data exchange and the commands and responses for the data exchange. The communication link control level performs information coding, data transfer control, error checking and error recovery, as well as some more complex functions. The message handling level performs message routing and flow control, message disassembly and reassembly, and other functions as might be needed for transmission methods such as message switching or packet switching.

Protocol considerations are very important for distributed computer networks because of the heavy reliance on communications. If single family computer hardware is used throughout the system, the networking task is not difficult, but if a variety of devices are used the task can be immensely complex. The subject of local networks will be covered in the next section of this report. Protocols are important in that they affect the flow of information throughout the total network. Consequently, a more detailed discussion of this topic follows later in this section.

COMMUNICATION NETWORKS

The selection of an appropriate communication network for a distributed computer system involves examining cost and performance features of viable alternatives. The performance goals of the communication network are generally broad in scope to allow for variation in the expected volume of data to be transmitted, to provide the desired system speed of response, and to provide the necessary level of system reliability and availability. The design goal for the communication system, then, is to achieve these performance features at minimum cost. Since many alternatives are presently available, and since new communication technologies are emerging, this task can be quite complex.

Three broad categories of communication services can be defined for purposes of this discussion: user owned (or private) transmission links, public common-carrier links, and leased (or private) transmission links. There are few large-scale, geographically dispersed communication links owned solely by the actual user. An exception is the United States military

establishment which operates radio and satellite networks. Some intermediate scale examples exist where nonmilitary users have their own microwave or radio links. On a local level, there are many user owned networks, with local area networks being a prime example. These local area networks are very important in distributed computer systems in general, and in CBI systems in particular. User owned links are the only truly private communications facilities, although leased communication lines are often referred to as "private."

A common carrier (such as the Bell Telephone Company) is a government regulated company providing communication services to the general public. Two types of services are available--switched (or public) lines and leased (sometimes called private) lines. Some confusion over the term "private line" comes about now since there are companies that lease a line from a telephone company and then offer a service to a number of users. Thus, while the line is private in the sense that it is not part of the public switched network, it is shared by several users. The public switched lines are dial-up lines going through a central switching exchange which makes the physical circuit connection. This physical connection exists only for the time duration of the call. In this way, the overall network is shared among many users from the general public.

In general, leased common-carrier links include dedicated telephone lines, broadband links provided by microwave, radio links, satellite links, and other high technology links such as laser and fiber optics. The leased (sometimes called private) aspect means that the user who contracts for the communication service is not expected to share the use of the facili-ties. Thus, these communication facilities are always available to the lessee, and a certain degree of privacy and security are guaranteed since the "public" does not have access to the link. Leased communications links are in widespread use by business, industry, universities, and government. These links are expensive, and their use must generally be justified on a cost versus performance basis.

In some cases it is more cost-effective to share a communication link or network with other users so that the entire cost is not borne by one user. Such shared services are offered by companies which lease lines or other facilities from a common carrier and typically add computerized switching and error detection/connection services. These public networks are sometimes called value-added networks. Two major examples are TELENET and TYMNET. TELENET began operation in 1975 and is based on the technology developed on the DOD ARPANET. Both TELENET and TYMNET are national in scope and offer coast-to-coast services in every contiguous state. Packet switching is typical of the value-added services offered by public networks, and it is a service offered by TELENET AND TYMNET. There are two significant advantages to a value-added public network. They are (1) use of high speed lines which might not be cost effective on a private basis and (2) provision for error detection to help ensure a high level of reliability.

Although TELENET and TYMNET are usually called public networks, they also offer private network services. Companies can arrange for their own

private networks by contracting with these or other firms offering similar services. In terms of evaluating alternatives for communications services, some combination of public network service and private network service may provide the most cost effective choice while still meeting the overall system design goals.

Other public network services have been offered for a number of years by Western Union Corporation's TELEX network and the American Telephone and Telegraph Company (AT&T) network. Newer companies offering public network services include Xerox Corporation with their Xerox Telecommunications Network (XTEN); a consortium of IBM, COMSAT General Corporation and Aetna Life Insurance Company with their Satellite Business Systems (SBS) network; and a new AT&T offering, Advanced Communications Service (ACS). There will be increased activity in this business area because of the Federal Communications Commission (FCC) ruling in 1980 to partially deregulate communications services to allow more competition.

The major independent common carriers other than AT&T include General Telephone and Electronics (GTE), Continental Telephone Company, United Telecommunications, and Western Union Corporation. In addition to these large, general purpose common-carriers, there are specialized common-carriers' who use microwave links to connect major facilities, and then use other carriers links to connect with customer locations. Some of these large, specialized common-carriers include Southern Pacific Communications Company (SPCC), United States Transmission Systems (USTS), and RCA American Communications, Incorporated (RCAACI). Table 3 summarizes the service and data rates provided by some of the common carriers and specialized common-carriers.

A special class of communications networks of some interest is a group providing what is often called "electronic mail" services. Three general types of networks exist: (1) general bulletin-board systems, (2) specialized bulletin-board systems, and (3) specialized information service systems. Bulletin-board systems are very numerous, and they are often free to the user. If users have the appropriate software and a modem, they can operate their own bulletin board or use other bulletin boards. In this way a message service is provided similar to a central bulletin board where people leave a variety of messages for others. The subjects of these messages are usually quite broad and may include items for specific people or people in general, such as items for sale or trade. Messages which might otherwise have gone to individuals by mail or to a newpaper or magazine could in many cases be appropriately routed to the bulletin-board systems.

The distinction between general and specialized bulletin-board systems relates to both the types of computers involved and the subjects of the messages. Specialized bulletin-board systems are a subset of the general system and might be restricted to certain computers such as APPLE or Radio Shack, or restricted to special interests such as music, science fiction literature, or military training. The third type of network, the specialized information service, usually is not free and provides data such as news, weather, financial reports, and other forms of current information. Some of the present network services available in this area are The Source, CompuServe, and Dow-Jones. These networks all charge fees for using their

services. In addition to a fixed monthly fee, there is usually a charge for the time a user is connected to the service. The dial-up, public telephone network is typically used to access the service, so this cost must also be added to the time-of-use fee. Thus, all cost components need to be considered when assessing the overall cost-effectiveness of a network service.

TABLE 3. COMMON CARRIER AND SPECIALIZED CARRIER SERVICE AND DATA RATES

Service Types/Transmission Speeds					
	Common Carrier	Specialized Common Carrier	Value Added Carrier	International Service	Digital Service
American Telephone & Telegraph (ATT) New York, NY	30-56,000 bps	—	—	—	2.4 Kbps-1.544 Mbps
General Telephone & Electronics (GTE) Stamford, CT	30-56,000 bps	—	75-56,000 bps	—	—
Western Union Upper Saddle River, NJ	75-50,000 bps	—	—	—	—
Tymnet Cupertino, CA	—	—	110-4800 bps	—	—
RCA Global Communications New York, NY	—	2400-9600 bps	—	12.5-56,000 bps	—
Southern Pacific (SP) Communications Burlingame, CA	—	75-56,000 bps	—	—	2.4 Kbps-4800 bps
MCI Communications Washington, DC	—	75-9600 bps	—	—	—
American Satellite Germantown, MD	—	—	—	—	9.6 Kbps-1.544 Mbps
ITT World Communications New York, NY	—	bandwidths to 240 Khz	—	12.5-56,000 bps	—
Western Union International New York, NY	—	—	—	12.5-56,000 bps	—

NOTE: All of the companies above provide a variety of communications services. Transmission speeds listed are a range encompassing all of the services provided.

DATA SECURITY

Data security can be defined as the protection of data from accidental or intentional disclosure to unauthorized persons and from unauthorized modifications. In a computer environment, deliberate or accidental disclosure of classified or otherwide sensitive data can take place even though commonly accepted data security countermeasures have been employed. The major reasons that data security is a potential problem in a computer environment stem from the computer itself and the manner in which it is customarily used. In an on-line operating environment, the resources of a computer system are normally shared among users to make effective use of the equipment. As a result, the special deployment of hardware and software facilities is necessary to provide the required level of data security. Modern distributed systems are particularly susceptible to security problems because of the large amount of data transmitted by telecommunication systems. The data communications system normally carries current information which is typically the most sensitive and valuable. Consequently, data privacy and security considerations are extremely important.

The need for data privacy and data security is important in many applications. Security may be desired for classified government data or for confidential business data. CBI systems dealing with classified military data will have special requirements for security. Management information systems for business and industry are vulnerable because of the processing, storage, and communication of company confidential records pertaining to sales, profits, financial projections, pricing policies, and market analyses. Data privacy is also of concern when personal information such as births, deaths, health status, income, taxes, and welfare payments are involved. In a CBI system, privacy of student performance data and personnel data is important.

The subject of data security has been explored by many users of distributed computer systems including Federal and state government agencies, computer hardware manufacturers and computer software producers. While the communications system is the most vulnerable part of a computer-communications system, there are other parts which are also vulnerable. Data security measures include an organized set of procedural, hardware, and software facilities that collectively prevent an unauthorized person from obtaining information from a computerized data management system. Available data security measures can be grouped into the following six broad classes:

- access management -- also called access control or user identification
- processing limitations -- such as memory protection and procedural restrictions on copying or modifying files
- auditing and threat monitoring -- where attempts to violate the security of a computer system or data file are recorded

- privacy transformations -- where the use of cryptographic techniques are used to conceal the contents of a message, record or data file

- integrity management -- where the physical security is maintained by the use of locks, controls, safes, restricted admission, and other similar measures

- level of authorization and data file protection -- where a given user is limited to his own data files and programs, and shared program use is restricted and recorded.

The basic technique for protecting data sent over a communication link in a system is that of encryption. This technique encodes or alters the data being transmitted so that its information content is concealed from the casual observer. An authorized receiver can properly decode the data for its intended use. The process of encoding and decoding data is generally referred to as cryptography, although some commercial computer system people call it privacy transformation (Katzan, 1979). Military and government users are highly skilled in encryption techniques and, in general, use much more complex codes and methods than businesses and industrial firms.

Encryption may be defined as a reversible set of logical and arithmetic operations that are performed on the characters of a message or a data record to make the information unintelligible for computer processing or human recognition (Katzan, 1979). In practice, a message is encoded by means of a key (a set of characters), and that same key must be possessed by the receiver of the new data in order to decode it. Encoding of the original data can occur off-line or on-line. The off-line approach involves preprocessing by some device before the data is input to the actual computer system for transmission. The on-line approach requires the use of the computer itself, or a special device, to encode the original data before sending it out over the communication link. Thus, encoding and decoding can be performed by either hardware or software, but additional hardware is usually preferred in order to protect the coding algorithm from discovery by some user of the computer system. Integrated circuit chips are now available to perform this task.

Commercial users of encryption usually adopt the standard techniques developed by IBM and formalized by the National Bureau of Standards (NBS). Except for military and diplomatic services, the NBS techniques are becoming the standard for all systems (Booth, 1981). This NBS approach involves encoding and decoding blocks of data consisting of 64 bits under the control of a 64-bit key. The decoding process is the reverse of the encoding process and must make use of the correct key. Since the NBS algorithm is widely known, the success of the system depends on keeping the keys protected, and this causes another type of security problem.

In a distributed computer network, there may be a need to perform multilevel encoding of data, depending on which node in the system is the

intended receiver. Thus, if node A sends a message to be understood only by node C, but must first pass through node B as part of the communication link, then some form of encoding must be used which will defeat the decoding capability of node B. This approach may be particularly important if a public network is used where the data may be seen by many legitimate as well as nonlegitimate receivers as it passes through the network to its final destination.

For Navy CBI systems, a basic issue will be one of data classification and classified data containment. If classified data is considered to be that data which could affect national security then effective data protection is essential. If there is no classified data then the concern is one of data privacy. Here the expectation is that the value of the data is not sufficiently great to warrant extreme protective measures. Rather simple encryption techniques can be used successfully to protect personal data. However, if classified data is involved, there must be extreme protective measures applied to comply with security requirements. In general, there is a trade-off between the degree of security and privacy desired and the cost of providing such security and privacy. It is not cost effective to highly protect data that is of little value. On the other hand, classified data must be highly protected even though the cost of protection might be significant.

PROTOCOLS AND INTERFACE STANDARDS

A protocol is a set of conventions or rules for defining information interchange procedures in a computer network. Although the terms "protocol" and "interface" are often used synonomously, protocol is really a more general term encompassing "interface" and other factors. The ISO reference model for protocols involves seven levels or layers and is used to establish an overall framework to enable systems to be inter-connected and to communicate with each other. This ISO reference model is shown in figure 7. Level 1 is the physical connection layer which specifies the electrical and mechanical hardware standards. An example of this level is the common RS-232 standard for serial communication between devices. Level 2 is the data link layer which specifies the rules for transferring data and control information over the communications link. An example of this level is high-level data link control (HDLC), another protocol defined by ISO. Another example is IBM's synchronous data-link control (SDLC).

Level 3 involves message switching and routing determinations over the network, and specifies the end-to-end logical conditions between system processes. There is a proposed standard for level 3 from the Institute for Information Processing (IFIP), called the IFIP internetwork end-to-end transport protocol. The X.25 protocol from the International Consultive Committee on Telephone and Telegraphy (CCITT) does include some end-to-end requirements, but these are ill-defined or incomplete (Weitzman, 1980). Level 4 is the user protocol layer and might actually consist of several protocols for separately handling files, terminals and remote jobs. At the present time, level 3 and 4 protocols are generally written for a particular computer or family of computers and, therefore, are unique to that system. DECNET is such an example.

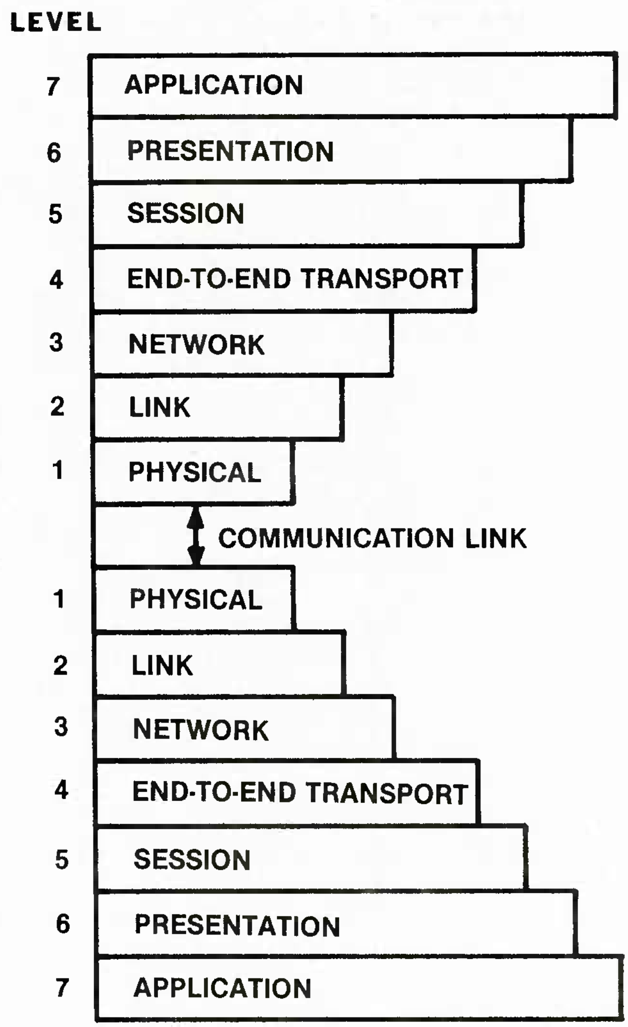


Figure 7. ISO Reference Model

The X.25 protocol is important in its own right as it is the ISO standard for international data communications. It is an interface standard for public packet-switching networks like TELENET, TYMNET AND UNINET. A particular feature of interest is that X.25 provides a bridge to the IBM SNA (discussed below). The X.25 standard protocol is used widely in Europe and Canada, thus allowing connection through worldwide networks. Local networks of microcomputers can also be connected to distant geographic points through an X.25 gateway into the packet switching networks.

Another important protocol-related topic is the IBM System Network Architecture, or SNA. IBM has a large share of the data processing market and is heavily involved in supporting applications involving network processing, or "teleprocessing" in their terminology. Because of the large number of IBM system users, the SNA is a de facto standard and must be considered in terms of interconnecting a non-IBM system to an IBM system. SNA is IBM's own internal standard and replaces a wide range of previous protocols and systems that they had developed over a period of years. SNA aims at providing an integrating network structure that is broad enough to satisfy the diverse requirements of various communications system configurations, and flexible enough to be adaptable to particular dedicated applications (McGlynn, 1978). Many computer systems from other manufacturers offer SNA compatibility in order to take advantage of this large market. SNA contains several protocols, but the major one is SDLC to handle the data link layer. In addition, as mentioned above, IBM has recently offered an X.25 capability with SNA so that IBM systems can make use of the public packet switching network.

THE X.25 INTERFACE STANDARD

Several references have been made to the X.25 standard from the International Consultive Committee on Telephone and Telegraphy (CCITT) in terms of a protocol and as a uniform access method for packet switching networks. This subsection provides a greater level of detail for this important communication standard. Although X.25 originated as a uniform access to packet switched networks, it is now serving as a common link for communications between computer products from many different manufacturers. The goal of many computer network planners is to have an "open systems" environment where any device can communicate directly with any other device.

As recommended by CCITT, X.25 has three levels or layers--physical/electrical control, communication link control, and packet control. These levels are shown in figure 8.

Level 1 of X.25 is the physical/electrical interface which defines the modem connections, which are usually RS-232C. The CCITT standard X.21 is an interface standard which can also be used for this level. Level 2 of X.25 manages the communication link interface, such as that needed to link a node processor to the actual communication network. The protocol used for this level is HDLC (high-level data link control) which is a bit oriented protocol that is also an ISO standard. This communication link control protocol describes the format of the data to be transmitted, which is called a

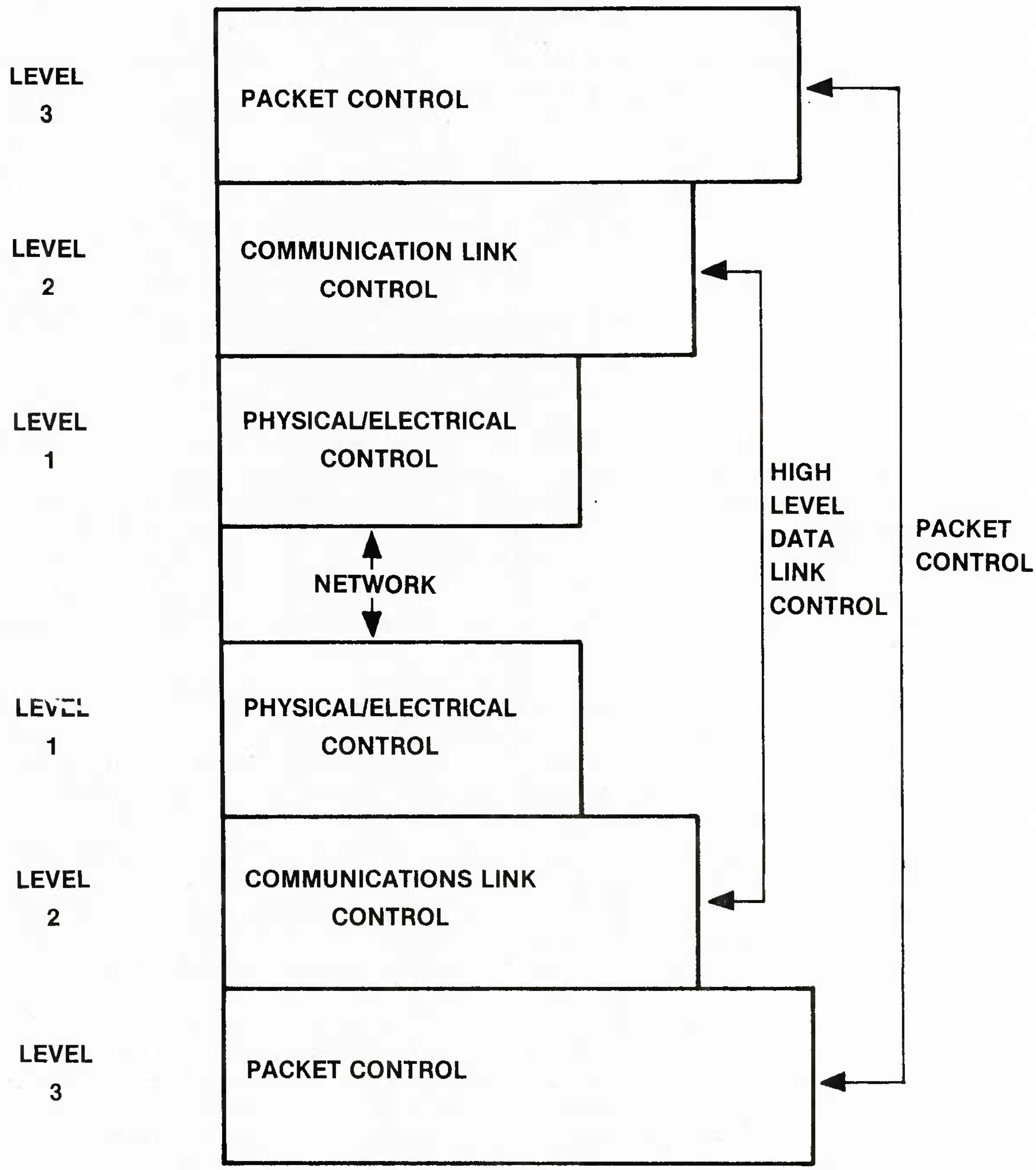


Figure 8. Protocol Levels For X.25

frame. This frame format is shown in figure 9.

DATA TO BE TRANSMITTED

FLAG	ADDRESS	CONTROL	Information Field	ERROR	CONTROL	FLAG

Figure 9. X.25 Frame Format

The address field determines where the information is to be sent, and the control field contains status, command, and response information. There are three types of control field formats--the information format, the supervisory format, and the unnumbered format. The information format contains data which can be used to tell if any of the frames of the complete message were lost in transmission. The supervisory format contains status information, and the unnumbered format specifies data communication link control commands for which no feedback is needed.

Interfacing a host computer with a network using the X.25 protocol is based on the concept of a virtual circuit, which is a logical rather than physical connection between the points in a network. A given logical connection may actually specify a particular physical route through the network, but, in general, a great amount of flexibility in route determination is possible. Many virtual circuits can be established simultaneously by a node processor by assigning a logical channel number to each destination, and then using that number to identify the particular data being sent to the different destinations. The virtual circuit interface is currently offered by all the public packet switching networks such as TYMNET and TELENET.

The X.25 protocol is one of the most important protocols because of its use as a uniform access to packet switching networks and its acceptance as a general interface between computer products from different equipment manufacturers. There are two major considerations involving the use of X.25 in Navy CBI systems. The first is that it may be desirable to use a public packet switching network for some CBI application, and X.25 capability would have to be part of the system design. The second is that X.25 capability would allow a given CBI system to interface with many other Navy and non-Navy computer communication networks to exchange information of a diverse nature. In particular, an interconnection of CBI systems would allow for the exchange of valuable operating data related to courses of instruction, evaluation results, and new developments.

THE RS-232C STANDARD

The RS-232C standard for serial communication between user's Data Terminal Equipment (DTE-- such as a terminal) and Data Communications Equipment (DCE--such as a modem) is widely accepted in almost all applications of computer and communications systems. The Electronics Industries Association (EIA), is the group sponsoring the RS-232C standard. During the early computer development period, this group recognized the need to establish some common interface standard for connecting a number of devices from different manufacturers to a given computer system. Without such a standard, it would be very difficult for different manufacturers' equipment to communicate with each other. It should be again noted that the RS-232C standard deals only with the interface of a device to a communication line and does not specify how the actual communication line or process will work.

Essentially, RS-232C guarantees that:

- voltage and signal levels will be compatible
- the interface connectors may be plugged together mechanically (mated) and maintain correspondence between pin and wire connections
- certain control information supplied by one device will be understood by the other device.

In addition, it specifies the functional interchange circuit description for serial binary data exchange, two primary channels and two secondary backup channels, one in each direction. There is also a maximum distance limitation of 50 feet between devices. Some of the commonly used signals specified by the RS-232C standard are given in table 4 (Tektronix, 1978).

Not all devices which claim RS-232C compatibility will in fact work together since some manufacturers use signals in different ways and some use nonconnected pins for their own special purposes. A compatible RS-232C device must have both the correct pin connections and the correct information on those pins. Several considerations that are not specified by RS-232C that are essential from a user's point of view include clock timing pulses, some control characters, and the proper idle characters. There are also mechanical interface differences among some RS-232C cables, so not all equipment will connect together with a given RS-232C cable.

Some of the operational limitations of RS-232C include the restriction to 50 foot maximum distances between devices and the restriction to a maximum 20,000 bits per second data transfer rate. The 50 foot distance limitation can be removed by using line driver amplifiers, but this would be an added cost and might result in lower system reliability. The limitation to 20,000 bits per second information transmission speed is becoming a severe problem since data rate requirements are growing at an ever increasing rate. Another factor that limits the attractiveness of RS-232C is the lack of any capability to test the data communications associated with the devices.

TABLE 4. COMMONLY USED RS-232C SIGNALS

SIGNAL NAME	MNEMONIC	USE
Transmitted Data	TDATA	Outgoing data path from the terminal's point of view
Received Data	RDATA	Incoming data path from the terminal's point of view
Request to Send	RTS	Activated by the terminal to tell the modem to prepare to receive and retransmit data from the terminal
Clear to Send	CTS	Activated by the modem to tell the terminal that it is ready to receive and retransmit data from the terminal
Data Set Ready	DSR	Activated by the modem to tell the terminal that the modem is operational
Signal Ground	SG	Return path for all other signals on the bus
Received Line Signal Detector	CD	Activated by the modem to tell terminal that the modem has made contact with the modem at the far end, and can sense the carrier
Data Terminal Ready	DTR	Activated by the terminal to tell the modem that the terminal is operational

In order to update the RS-232C standard, as well as to upgrade it for presently envisioned tasks, several new standards have been adopted. The first new standard is RS-423A which covers the need for unbalanced-voltage, digital-interface circuits. The second new standard is RS-422 which covers balanced-voltage interface circuits similar to the RS-232C for high speed binary data exchange. This results in better noise immunity and lower cross talk, which results in lower transmission error rates. The new standards also provide a substantial improvement over RS-232C in terms of safety for integrated circuits which use low voltages. The 25 volts on RS-232C pins often results in chip destruction when connections are made improperly.

A new general standard which incorporates RS-422 and RS-423A is RS-449, which will most likely replace the RS-232C. Some of the features of RS-449 include data rates up to 10 million bits per second and cable lengths up to 4,000 feet. This RS-449 standard should be coming into widespread use in the future, especially since the Federal government has mandated that all procurements after June 1, 1980 be RS-449 compatible (Sherman, 1982).

ARPANET

The ARPANET system has been mentioned several times in this report as an example of a communication network. This network is important since it is a major interconnection of computers at different contract research centers of the Advanced Research Projects Agency (ARPA) of the Department of Defense. In terms of computerized instruction systems for Navy training, the ARPANET facilities are potential candidates for data transmission and processing. ARPANET was started as an experimental project in 1969 to study the large-scale interconnection of different types of computers that were widely separated geographically. It was not the first computer network, but because of its size and scope it has received a great deal of attention. The main purpose of ARPANET is resource sharing, allowing research workers at different locations, including many universities, to exchange information and make use of each other's unique resources (Booth, 1981). ARPANET is generally considered the parent of all modern computer networks in general, and all packet switching networks in particular.

The ARPANET is a collection of large host computers (over 100 in 1979) interconnected through an independent packet switched communication system composed of Interface Message Processors (IMPs) and Terminal Interface Processors (TIPs) which are primarily Honeywell 316 and 516 microcomputers. A geographic map of this network is shown in figure 10, and a logical map is shown in figure 11. By 1979, Honeywell had stopped manufacturing the 316 and 516 minicomputers so newer IMPs and TIPs use other processors. Bolt, Beranek and Newman, Incorporated (BBN) produces IMPs and TIPs that have much greater capacities and throughput than the older Honeywell units. The BBN devices are called Pluribus IMPs and Pluribus TIPs.

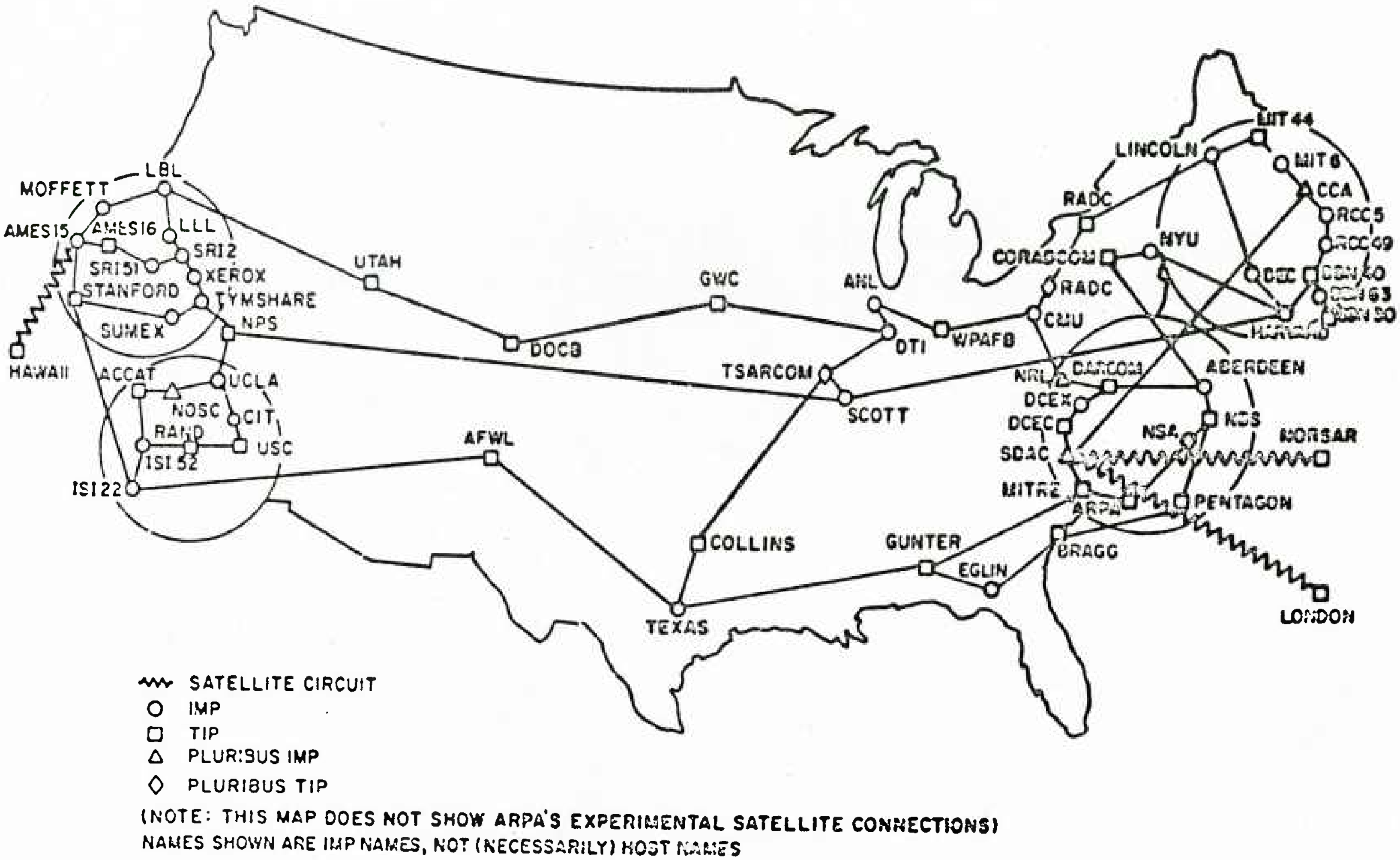


Figure 10. ARPANET Geographic Map, September 1979

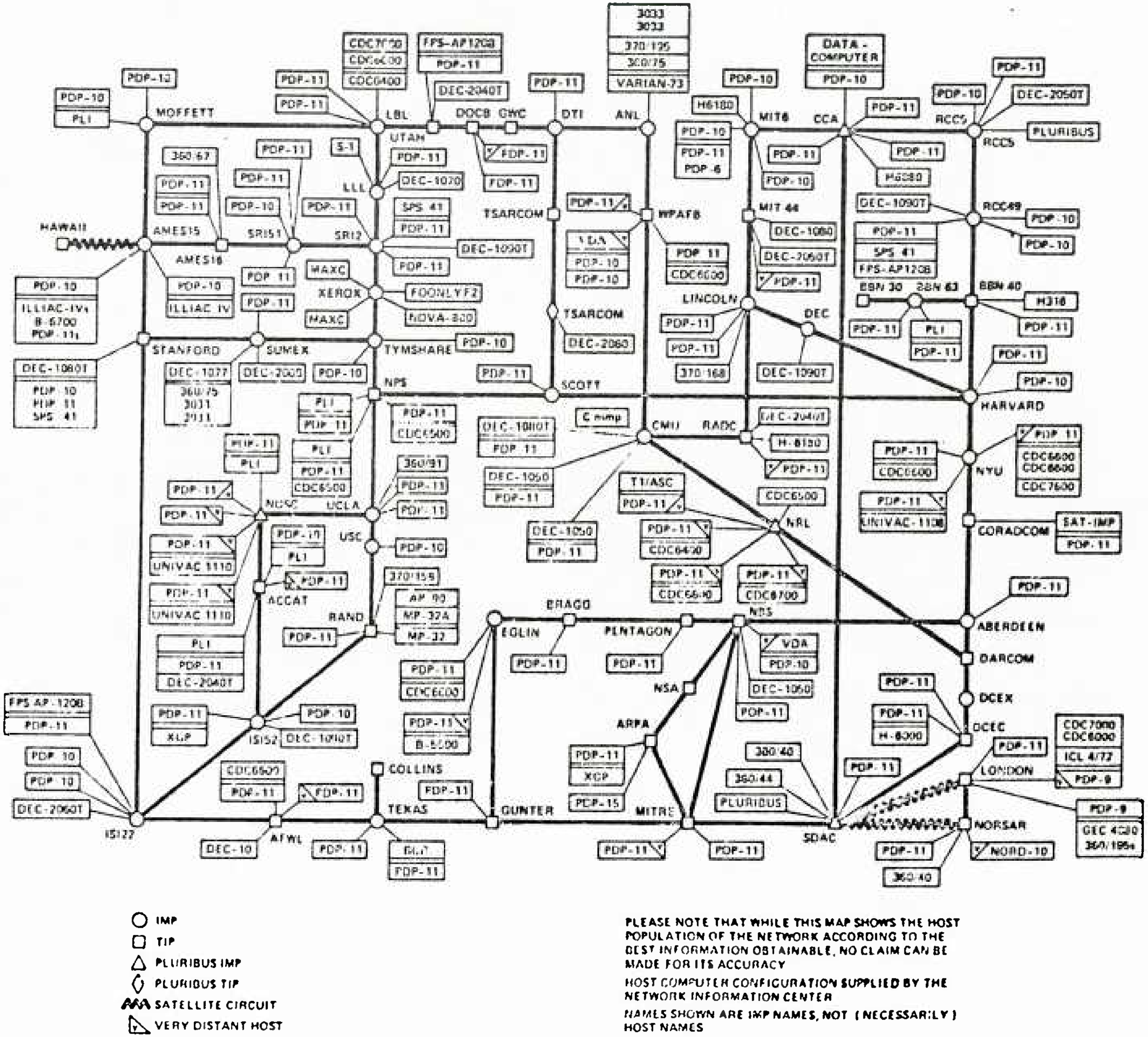


Figure 11. ARPANET Logical Map, June 1979

A brief description of the ARPANET is given in the ARPANET Directory produced by the ARPANET Network Information Center at Stanford Research Institute, Menlo Park, California. This description is reproduced below, as well as a listing of who is eligible to participate as a user of ARPANET.

BRIEF DESCRIPTION OF THE ARPANET

The ARPANET is an operational, computerized, packet switching DoD digital network which provides a capability for terminals or geographically separated computers, called hosts, to communicate with each other. The host computers often differ from one another in type, speed, word length, operating system, and other characteristics. Each terminal or host computer is connected into the network through a small local node computer called an IMP or TIP. The complete network is formed by interconnecting the IMPs through wideband communication lines (normally 50,000 bits per second) supplied by common carriers.

Each node is programmed to receive and forward messages to the neighboring nodes in the network. During a typical operation, a host passes a message to its node; the message is passed from node to node through the network until it finally arrives at the destination IMP, which in turn passes it along to the destination host. This process normally takes less than 250 milliseconds.

Hosts communicate with each other via regular messages. A regular message may vary in length from 96 to 8159 bits, the first 96 of which are control bits called the leader. The leader is also used for sending control messages between the host and its IMP or TIP (node). The remainder of the message is the data or text.

For each regular message, the host specifies a destination, consisting of node, host, and handling type. These three parameters uniquely specify a connection between source and destination hosts. The handling type gives the connection specific characteristics, such as priority or non-priority transmission. Additional leader space has been reserved for a fourth parameter, to be used in future internetwork addressing. For each connection, messages are delivered to the destination in the same order that they were transmitted by the source.

For each regular message, the host also specifies a 12-bit identifier, the message - ID. The message - ID, together with the destination of the message, is used as the "name" of the message. The node uses this name to

inform the host of the disposition of the message. Therefore, if the host refrains from re-using a particular message - ID value (to a given destination) until the node has responded about that message - ID, messages will remain uniquely identified and the host can retransmit them in the event of a failure within the network.

After receiving a regular message from a host connected to it, a node breaks the message into several packets (currently the maximum data bits per packet is 1008) and passes these through the network in the direction of the destination. Eventually, when all packets arrive at the destination, they are reassembled to form the original message which is passed to the destination host. The destination node returns a positive acknowledgment for receipt of the message to the source host. This acknowledgment is called a Ready for Next Message (RFNM) and identifies the message being acknowledged by name. In some relatively rare cases, however, the message may not be delivered due to a node failure; line disruption, etc., in such cases an Incomplete Transmission message will be returned to the source host instead of a RFNM. In this case the message which was incompletely transmitted is also identified by name.

If a response from the destination node (either RFNM or Incomplete Transmission) is not delivered to the originating host, this condition will be detected by the source node, which will automatically inquire of the destination node whether the original message was correctly received and repeat the inquiry until a response is received from the destination node. This inquiry mechanism is timeout-driven, and each timeout period may vary between 30 and 45 seconds in length.

When a message arrives at its destination node, the leader is modified to indicate the source host, but the message - ID field is passed through unchanged. Thus, in addition to providing message identification between a host and its local node, the message - ID can provide a means for hosts to identify messages between themselves.

Users of the ARPANET may access local or distant SERVER computers (hosts) over the network. They may also exchange messages, create realtime links between users, transfer files from one computer to another, and submit batch jobs to distant computers. For a more complete description of these processes, see the ARPANET Protocol Handbook available from the NIC or the National Technical Information System (NTIS), Springfield, VA. 22161 as AD A052594.

WHO MAY BE A SUBSCRIBER ON THE ARPANET

DOD USERS

Subject to the availability of assets, DoD activities will be connected to the ARPANET provided the requests are processed through normal communications validating channels.

NON-DOD U.S. GOVERNMENT ACTIVITIES

Requests for ARPANET service from non-DoD U.S. Government activities will be considered by DCA on a case-by-case basis.

NON-GOVERNMENT U.S. ACTIVITIES

A DoD or other U.S. Government activity authorized to use the network may sponsor, as a user, a non-government activity performing contract work for the U.S. Government. Justification outlining benefits to the U.S. Government for such access shall be provided to DCA by the sponsoring activity. Cost for network services provided to non-government activities will be allocated to the sponsoring activity.

SECTION VI

LOCAL AREA NETWORKS

This section addresses a number of the decisionmaking considerations relating to the use of local area networks in typical Naval training settings. These specialized networks are also of importance for general commercial, business, and industry applications. Local area networks, or LANs, are one of the rapidly developing areas of computer system technology, with the present 5,700 applications expected to triple to 20,000 by 1986 (Small Systems World, 1982). Computer system engineers and consultants suggest that any user considering the use of a minicomputer in a training application should also consider the option of a microcomputer network to perform the same task. Some of the advantages to be realized with a microcomputer network would be high reliability, rapid response, and structured resource sharing. Many users are currently replacing collections of telephone lines and other low speed communication links with local area networks which are functionally less complex and considerably more adaptable to changing requirements.

A local area network is a distributed processing system confined to a local area of a few square miles and as such offers the advantages and disadvantages discussed previously. For applications where geographically close elements in a distributed processing system are the design of choice, a LAN will probably be a necessary addition to the hardware and software components. LANs will be particularly useful to two categories of users. The first is the group of users who already have computer networks which are approaching the limits of their current capabilities. For this group LANs offer an effective, low cost method to add network capabilities while extending geographic limitations since LANs can be connected together through long haul networks, or through satellites. The second group of users consists of those who are not now networked but who desire to distribute their processing functions for reasons of efficiency and effectiveness. Navy CBI systems fall into this second category. Integrated business office applications and manufacturing operations--especially computer aided design (CAD)/computer aided manufacturing (CAM)--also fall into this category.

In general, local area networks have the following basic properties:

- controlled by a single organization
- geographically limited; that is, a local network's backbone spans a distance on the order of only a few miles
- contain some type of switching technology
- transmission rates are usually faster than those of networks covering a broad geographic area.

Some of the well known commercially produced LANs are ETHERNET from Xerox Corporation, WANGNET from Wang Laboratories, OMNINET from Corvus Systems, Cluster/One from Nestar Systems, NET/ONE from Ungermann, Bass Incorporated and SERIES/1 RING from IBM. The Navy currently uses a local network called Local Net at the Surface Weapons Center in Dahlgren, Virginia, which is produced by SYTEK, Incorporated.

LANs are, in essence, an example of a complete distributed network with the general limitation of network nodes being located within a local area. Within the framework of local area networks, the two most important general topics of concern are network architecture and communications technology. Several aspects of communications technology will be discussed, including transmission mediums, data rates, and network software for communications processing and network switching. These topics and other related topics will be addressed in this section. A subsection on the application of LANs to Navy CBI systems is also included as an applications oriented summary for this section.

TRANSMISSION MEDIUMS AND DATA RATES

LANs can be based on any of several communications technologies depending on the information needs of the distributed system and the data transmission rates required. If a private branch exchange (PBX) or computerized private branch exchange (CPBX), both of which are communications networks, is available in an office, a factory or a military training center, then it is possible to use this network to support the distributed system. If an existing communications network is not available, then some transmission medium such as wire cable, coaxial cable or fiber optics cable must be installed to support the system. The form of information to be communicated and the required data rates are usually the factors which provide a requirements baseline for specifying the transmission medium needed.

Essentially four types of information can be transmitted over a local network. These include digital data which is exchanged between system nodes, audio data such as telephone conversations (which can be modulated digital data), video signals such as television, and image signals used in applications such as CAT scanners and industrial CAD/CAM operations. At this point, only the information flow related to computers or terminal devices has been addressed. Although this is probably the most significant aspect of information flow in a distributed processing system, it is important to consider the uses of the other forms of information which can be transmitted over a local area network. By using the full capability of a LAN, the cost effectiveness can be improved for the entire system.

The forms of information transmitted over a LAN and the associated data rates depend on whether the network is of a baseband or broadband design. These terms refer to the frequencies of transmission where baseband transmission is single channel while broadband transmission can contain many channels simultaneously. The use of a single twisted pair wire or a multiconductor cable is restricted to baseband operations, and the maximum data rate is about 1 megabit per second. The advantage of this medium is that it is relatively low cost, but it has serious disadvantages--lack of physical ruggedness and susceptibility to electrical noise pickup. The low data rate capability means that while audio and digital data can be transmitted over such wires, these wires cannot be used to transmit video data or high speed digital image data. If those latter two features are not necessary for a given system design, then the wire approach may be very desirable. Coaxial cable can also be used in a baseband mode and offers a different set of advantages and disadvantages. It is more costly than twisted pair wire, but it is also

much more physically rugged. It can be installed easily without fear of damage, and coaxial cable taps are readily available. The speed of transmissions along a coaxial cable can be increased to about 10 megabits per second, which is a significant gain over the twisted pair wire. This still does not allow video or image data, though, if other services are combined on the one cable.

The use of a broadband transmission mode requires a radio frequency (RF) modulator to achieve the wide bandwidths and the subsequent high speeds of data transmission. Coaxial cable is usually the medium preferred in broadband LANs, but fiber optic cables may also be used. Unshielded, twisted pair wires or cables are unacceptable mediums for broadband transmission. The speed of transmission along a broadband coaxial cable is usually limited to 50 - 100 megabits per second although it may go as high as 250 megabits per second. This high speed capability is quite often partitioned into a series of lower speed channels instead of using the entire bandwidth for one channel. For example, a video TV transmission requires only the equivalent bandwidth of a 6 megabit per second rate, so if a 50 megabit per second capacity is available then several services can be provided at one time over the same cable. The operation of a broadband LAN requires that all signals on the communication system be converted to RF signals, and then transmitted over the coaxial cable on separate channels. Thus, a single broadband coaxial cable can replace an entire network of twisted pair wires or telephone cables and probably still have spare capacity for future expansion.

These broadband RF transmission systems are expensive but they have the advantage that the signal attenuation along the coaxial cable is quite low, so signals can be sent for several miles along the cable by highpower transmitters rather than several thousand feet as with baseband mode. The cost of the RF coupler or RF modem units for broadband LANs is about 10 times that of the low speed baseband coupling units because of their wide bandwidths and the need to separate out the various channels of information being transmitted. Since each device connected to a broadband cable must receive and transmit data, there is a need for retransmitters in the system if very great distances are involved. While a powerful transmitter such as a cable TV station may be able to provide signals over many miles, the low power RF modems cannot match this performance. Thus, repeater elements must be used in a typical LAN to provide communication over longer distances in a large building or between building complexes.

The cost of a connecting interface device for a node of a LAN is greatly dependent on the speed of the data transmission involved. Baseband networks using twisted pair wires have the least expensive connection, with several such commercial LANs having costs as low as $500 per node connection (Data Communications, 1981). One LAN with the capability to serve 30 microcomputer nodes costs only $100 per connection. Baseband LANs using coaxial cable are typically more expensive, having costs of $500 to $1,500 per node connection. Finally, broadband LANs have the highest cost per connection if the high speed capability of the network is utilized. If data rates of only 10 megabits per second or lower are used on broadband then the costs about match that for baseband coaxial cable systems. One commercial

LAN that has a data rate of 50 megabits per second--a truly wideband, fast system--costs $39,000 per node connection. Thus, it is clear that the cost of excess capability in a system is very high.

LOCAL AREA NETWORK ARCHITECTURE

Since LANs are examples of general distributed processing systems, the earlier section on network architecture applies directly to their design and operation. The six generic architectures previously discussed--star, ring/loop, multidrop tree, multidrop bus, irregular mesh, and fully interconnected mesh--could be used for LANs. However, three of these structures are not particularly appropriate for connecting a group of computers or devices that will be separated by distances of typically several hundred to several thousand feet. The multidrop tree is a hierarchical structure, and most LANs are either not hierarchical systems to begin with or are part of a hierarchical system whose nodes are on a uniform level--usually the lowest--of a parent hierarchical system. The irregular mesh structure is unlikely to be used in an LAN that is relatively compact in terms of distances between nodes. This particular structure has more value in networks where very great distances between nodes are involved. The fully interconnected mesh is also very unlikely to be used in a LAN for the opposite reason given for the irregular mesh--that the distances between nodes are too great. This structure is most applicable to the interconnection of several processors and peripherals which are physically close--such as in one large room--where short bus cables can be run from each device to every other device. This is already a costly structure, and the added costs of even a few more feet between system nodes would quickly escalate to unreasonable values.

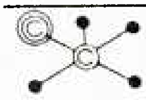
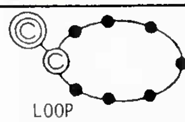
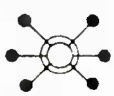
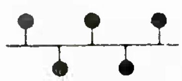
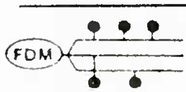
Three of the generic network structures are highly applicable to local networks--the star, the ring/loop, and the multidrop bus. The star structure already exists in many applications because of the use of the PBX or CPBX system. Here, a central facility is needed to handle the control node functions of switching and processing. For applications where data transmission volumes are low, these PBX/CPBX systems may be very cost effective for users that already have them. The ring/loop structure is presently one of the most prevalent for LANs. Some network designers draw a distinction between a ring and a loop, where a loop is most often used to connect a series of terminals together while a ring usually connects a series of computers. Others draw no distinction between rings and loops and consider them as virtually identical in architecture. The bus structure is probably the most common network structure for LANs, and can be used for either baseband or broadband transmission.

These three architectures most applicable to LANs are further described in table 5. The table draws a distinction between rings and loops, and also distinguishes between baseband bus systems and broadband bus systems.

The star network has a central controller which directs all communication between nodes of the network. As such, all star networks are vulnerable to a failure in this central control computer. A central control failure can disable the entire network, so it is very important to have a

highly reliable facility at this hub of the system. One benefit of the dependence on the central node is that the star network is highly secure from user device malfunction. If a failure occurs at a noncentral node, or in a communication link, there is no impact on the operation of the remainder of the system. Of course the data the failed node was providing is no longer available, but the basic equipment operation of the other nodes is unaffected. In general, star networks have an excellent record of reliability (Levy and Mehl, 1982). Since many star LANs will make use of existing PBX/CPBX telephone systems, they can be implemented easily, effectively, and at reasonably low cost. A diagram of part of such a star/CPBX system is shown in figure 12.

TABLE 5. LAN ARCHITECTURE

TOPOLOGY	TRANSMISSION MODE	TYPICAL PROTOCOLS	TYPICAL NO. OF NODES	ADVANTAGES
 STAR-PABX	POINT-TO-POINT VIA CIRCUIT SWITCH OR COMPUTER MEMORY	RS-232C OR COMPUTER	TEN TO ONE HUNDRED	WELL-KNOWN LARGE BASE OF USERS
 LOOP	MESSAGE ROUTING VIA LOOP CONTROLLER	SDLC	TENS	WELL-KNOWN LARGE BASE OF USERS
 RING	PACKET TRANSMISSION AROUND RINGS	HDLC (TOKEN PASSING)	TENS TO HUNDREDS PER CHANNEL	DISTRIBUTED CONTROL NO CONTENTION POPULAR FOR COMPUTER NETS
 BASEBAND BUS	BROADCAST ALONG SERIAL BUS	CSMA/CD	TENS TO HUNDREDS PER SEGMENT	DISTRIBUTED CONTROL POPULAR FOR OFFICE NETWORKS & COMPUTER NETS
 OTHER SERVICES BROADBAND BUS	PACKET BROADCAST BUS WITH DEDICATED CHANNELS	CSMA/CD RS-232C & OTHERS PER CHANNEL	TWO TO HUNDREDS PER CHANNEL	DISTRIBUTED CONTROL LARGE VARIETY OF USERS AND SERVICES

Reprinted with permission from Small Systems World, June 1982

Two types of star/PBX LANs are offered by commercial vendors: the integrated voice/data switch systems and port contention systems. The integrated voice/data systems use microprocessor based telephone sets and incorporate analog to digital conversion. Port contention systems are designed to work with a time sharing computer and permit any of a number of users to access any computer port by dial-up methods. Integrated voice/data systems

additions are offered by Anderson-Jacobson (10XX), Inte Com (IBX), Lexar (LBX), Mitel (SX-2000), Datapoint (ISX), GET (1000/4600), Northern Telecom (SL-1) and Rolm (CEX) (Levy and Mehl, 1982). Costs per node for the added systems range from about $250 to $350 for newer type PBX systems. Port contention systems are offered by Develcon, DCA (sys 355), INFOTRON (TLA50), Gandalf (PACX), Micom (Micro 6000 and Micro 650). A typical cost for one of these systems is $30,000, plus $100 per node connection, with the ability to connect up to 500 nodes.

Figure 12. Star/CPBX Network

The star network is an attractive LAN independent of the PBX/CPBX system. Other network structures, such as rings/loops, are often constructed of sections that are individual stars. While few commercial LANs are advertised as star configurations, a multidrop bus network with a centralized bus control scheme is actually a star itself.

The ring/loop network has one major problem -- its high vulnerability to a node interface device malfunction. Since all information must pass through each node interface device, the failure of any one node facility may disable the entire network. The problem of dealing with hardware failure implications in a token passing ring/loop will be discussed next in this section. Nonetheless, ring/loop LANs are common, with some commercial offerings being Primenet, Stratalink, and IBM Series/1 Ring. Costs of node connection interface devices depend on whether the LAN is baseband or broadband mode, and the data rate. A typical 2 megabit per second node connection is around $4,000 to $5,000.

The multidrop bus network also suffers from the same problem as the ring/loop network; that is, a high vulnerability to a node interface or bus interface device malfunction. In bus systems, all user access devices are permanently connected to the bus and can transmit or receive at any time. If a bus interface unit fails and transmits an unwanted signal on the bus, the entire network will be disabled until the failed unit is located and repaired or removed. Alternately, if a bus interface unit fails and produces a short circuit on the bus, then this also disables the network. There are many advantages of the coaxial cable bus networks, and some experts have predicted that this technology will be the widespread choice for general purpose LANs in the future (Krutsch, 1981).

Multidrop bus LANs are offered by a number of commercial vendors and include Omninet from Corvus Systems, Net/One from Ungerman-Bass, Ethernet from Xerox Corporation and Wangnet from Wang Laboratories. Costs of node connections for bus LANs depend on whether the network is baseband or broadband and the speed of data transmission. Node connections for a 2 megabit per second bus LAN range from $500 to $1500.

LOCAL AREA NETWORK ACCESS

While baseband mode LANs can allow only one channel of information to be transmitted at one time, this single channel can be shared by many different users. When several users will share such a channel it is necessary to either multiplex the data passing through it or to otherwise control the access to the channel. A common method for time sharing a single communication channel is time-division multiplexing, referred to as TDM. However, since the single channel bandwidth must be time shared, the maximum speed of any one user in a TDM system will be restricted to a value well below that of the maximum speed of the total channel. In this approach to time sharing a single channel, each transmitter and receiver has one or more dedicated or nondedicated time slots. A TDM system using nondedicated time slots will require some form of centralized channel control. The use of dedicated time slots is somewhat easier, and allows the channel control to be distributed throughout the network, thereby increasing the system reliability.

The TDM approach has the advantage of simplicity in that specific time slots are allocated for various users. This means that a standard multiplexor which combines a series of low speed data streams into one high speed data stream could be used in some systems. If dedicated time slots are assigned to particular node interface units, then these devices must transmit only in their specific time bands and not in other time bands and, thus, have only a part of the entire channel bandwidth. This approach limits the data rate for each node to some portion of the channel, and no single user can ever have the full bandwidth for a burst of high speed transmission. Another limitation of this TDM approach is that if a user does not require their share of the channel capacity during some period of time, then that portion of the channel capacity is simply wasted while other users are forced to slow down their transmissions. This inefficiency is the trade-off result for having a simpler system of time sharing operations.

Another approach to time sharing a single channel is to control the access to the channel, allowing a single user to take the total bandwidth for a small fraction of time. Two such approaches are contention controlled access and token controlled access. A basic contention based access method is carrier-sense multiple access with collision detection, referred to as CSMA/CD. With CSMA, each interface unit first listens to the network to see if any data is presently being transmitted. If data is being transmitted, the interface unit waits until the channel is clear before starting its transmission. Once the channel is clear, the node interface unit begins transmitting. Sometimes, however, two or more interface units may initiate transmission simultaneously. Since all transmitting nodes also listen, they

know that a collision of the several data streams has occurred. This collision detection is the CD portion of the CSMA/CD approach. When a collision is detected, all nodes that had transmitted immediately stop transmitting and wait a random period of time before attempting to transmit again. Each interface device generates its own random delay period, so there is a reduced likelihood of a collision on the next attempt to transmit. A priority system can be built into the system by providing certain node interface units with lower average delay times than others. If a collision occurs a second time for given messages being transmitted by several nodes, then the delay time is doubled before the next transmit attempt. If the delay interval finally reaches a large enough value, an error routine is initiated. The CSMA/CD scheme is most commonly used in coaxial cable multidrop bus LANs.

Most baseband local networks use contention access control, including Ethernet from Xerox, Net/One from Ungermann-Bass, and Hyperbus and Hyperchannel from Network Systems Corporation. The entire bandwidth of the channel is available to every node but with the potential for some delay after its initial effort to transmit. This approach is well suited to applications in which most devices using the network possess bursty (low duty cycle) data transmission characteristics. Some networks use only CSMA or CSMA with some priority scheme instead of CD. If only CSMA is used, then there is a possibility that an unstable state of communications operation can develop. The CSMA/CD approach is the most widely used contention system and is contained in LANs such as Ethernet, Cluster/One, Net/One, Wangnet, and Local Net.

Another access control approach is to use token passing, where a token is a message granting a given node interface unit the temporary but exclusive right to transmit. Token access, or token passing, is used mostly on ring/loop networks but it can also be used on multidrop bus networks. With token access, network control is passed from node interface unit to node interface unit in a specific known order. When a given node processor has control of the network it can transmit a message and then relinquish control by passing the control token to the next processing node. The token access control method ensures a relatively tight control over the network and completely prevents more than one node interface unit from transmitting at the same time. When a node interface unit detects a message for itself, it simply removes that message from the network.

The token access control scheme is a simple approach, but it takes a complex set of algorithms to handle the node linkages and to recover from network error conditions. One advantage of the token access scheme is the same as that for contention access control -- availability of the full bandwidth of the channel to each node in the network. Token access control is also very fair or equitable to each node in the network since there is a specific order for passing the token, minimizing the chance that one or more nodes might use up an unfair share of the network capacity. While token access control eliminates the inefficiency caused by data stream collisions in a contention access control scheme, there is overhead introduced by passing the control token from node to node. A node wanting the entire channel bandwidth in a token access control scheme must wait for the token to arrive

at its node before it starts to transmit, and thus the network is not as available on demand by a given node as is a network using contention control.

Some loss in network capacity exists in token passing since a number of nodes with no data to transmit may be given the control token before it arrives at a node requiring service. This overhead must be subtracted from the total channel capacity. A priority scheme can be added to a token access control approach where certain nodes are passed the token more often than others. Such a priority addition should reduce the overhead loss. Another critical problem in a token passing scheme is how to recover from a lost token, a damaged token, or the existence of multiple tokens. These problems can arise from various equipment failures of interface units, and they must be able to be corrected by the network control function. Error recovery mechanisms must be part of the token access control scheme in order to detect and correct these problems that would otherwise disable the network.

The use of token access control is mostly restricted to ring/loop network structures by present commercial vendors, but this restriction is not really necessary. Some of the common LANs using token passing are Modway from Gould, Incorporated, Primenet from Prime Computer, Incorporated, and Stratalink from Stratus Computer, Incorporated. The Modway network is a multidrop bus while the other two networks are rings.

Another access control approach that is used commonly in star networks based on PBX/CPBX technology is an interrupt and transmit scheme. Here, a node interface unit first initiates a request to send (RTS) a message along an interrupt line to the central processor which then establishes a connection to the desired receiving node. Then the transmitting node can send its data to the receiving node. Other nodes desiring to transmit to either of these two nodes that are presently in use must wait until the particular communication link becomes available. This approach to access control is called a matrix-switched system and it is usually restricted to applications with fairly low speeds of data transmission. Data rates can be as high as 56,000 bits per second once a communications link has been established. For users who already possess PBX/CPBX systems, it may be worthwhile to examine the advantages and disadvantages of this access scheme.

So far this discussion of time sharing a single communication channel has been tied to a baseband mode LAN. Obviously, this is a highly critical aspect of baseband systems involving a number of users. These same considerations exist in a broadband mode system, however, even though more than one communication channel can be contained simultaneously in such systems. Each channel of a broadband system can be time shared in exactly the same manner as has been discussed up to now for a baseband system. Thus, a broadband LAN may be used to transmit several channels of digital information for different categories of users as well as transmitting several channels of video information at the same time.

The method used to place many different channels of information on a broadband network is a multiplexing approach based on frequency division multiplexing, or FDM. Here, outgoing data is used to frequency modulate an RF carrier on a particular channel to which the intended receiver is tuned. The number of channels that can be contained on a broadband cable network is quite large, with the common CATV system cable having 55 video channels of 6 MHz bandwidth each. Thus, a broadband network is virtually unlimited in its capacity, although the cost goes up greatly as the capacity is expanded. A broadband network requires each computer device on the network to be coupled with a modem that converts RF signals to digital signals for both transmitting and receiving. These RF modems are generally much more expensive than the relatively simple transmitters and receivers used in baseband networks.

Several common broadband LANs are Local Net from Sytek, Incorporated and Wangnet. Some of the vendors for broadband LANs offer a combination of FDM and TDM so that several subnetworks can be used for the same cable system. These FDM/TDM systems are available from Local Net, Cablenet from Amdex Corporation, and Mitrenet from Mitre Corporation. With these systems, digital data, analog data, and video data can all coexist on the same cable.

COMMERCIALLY AVAILABLE LOCAL AREA NETWORKS

An excellent survey of LANs that are presently commercially available was prepared by Data Communications magazine. This survey was published under the title "Local Networks at a Glance" and appeared in the December 1981 issue of Data Communications. This survey is reproduced in table 6 and appears with their permission.

MULTIPLE SERVICES ON LOCAL AREA NETWORKS

Broadband LANs will most likely evolve into local, hybrid configurations, with multiple transmission technologies including low speed work stations for text and graphics plus wideband channels for high speed digital video communications. The use of a broadband LAN with FDM/TDM will allow all these services on a single coaxial cable. Such a network gives complete frequency coverage and allows handling all forms of computing, word processing, voice, data, facsimile and video conferencing. This type of network is more expensive than a baseband, single channel LAN but it offers a tremendous capability plus future growth potential. Many services can be provided by a high speed baseband LAN, but these systems cannot support video or image data. However, for many applications, a baseband approach may be satisfactory. Xerox, with its Ethernet LAN, has produced an office automation system called the Xerox 860 information processing system. This system interconnects a wide range of system components including the 8010 information system (a high powered work station for professionals and analysts), the 8030 series of network file servers, the 8044 laser elec-tronic printer, the 8071 network communications server (a small computer oriented work station), and a selection of electronic typewriters for very low cost work stations.

TABLE 6. LOCAL NETWORKS AT A GLANCE

	NETWORK NAME	MAXIMUM DISTANCE SPANNED[2]	TRANSMISSION MEDIUM	ACCESSING SCHEME	SPEED	MAXIMUM NUMBER OF CONNECTIONS	APPROXIMATE PRICE PER CONNECTION[3]	MAJOR APPLICATION	TYPE OF CONFIGURATION
AMDAX CORPORATION BOHEMIA, N.Y.	CABLENET	OVER 75 MILES	BROADBAND COAXIAL CABLE	TIME DIVISION MULTIPLE ACCESS AND FREQUENCY-DIVISION MULTIPLEXING (FDM)	7 OR 14 MBIT/S	16,000	$2,500 TO $3,000	GENERAL PURPOSE	BUS (DIFFERENT BANDS)
APOLLO COMPUTER INC CHELMSFORD, MASS.	DOMAIN	3,250 FEET	BASEBAND COAXIAL CABLE	TOKEN PASSING	10 MBIT/S	OVER 100	$39,000 (INCLUDES APOLLO COMPUTER)	SCIENTIFIC AND EDUCATION	RING
CORVUS SYSTEMS INC. SAN JOSE, CALIF.	OMNINET	4,000 FEET	TWISTED-PAIR WIRE	CARRIER-SENSE MULTIPLE ACCESS (CSMA)	1 MBIT/S	64 MICROCOMPUTERS AND VENDOR'S PERIPHERALS	$500 TO $1,000	OFFICE AUTOMATION AND EDUCATION	BUS
DATAPOINT CORPORATION SAN ANTONIO, TEX.	ATTACHED RESOURCE COMPUTER (ARC)	4 MILES	COAXIAL CABLE	TOKEN PASSING	2.5 MBIT/S	255	NA[4]	GENERAL PURPOSE	BUS
DIGITAL EQUIPMENT CORPORATION MAYNARD, MASS.	DECDATAWAY	15,000 FEET	TWISTED-PAIR WIRE	HDLC-LIKE "BIT STUFFING" PROTOCOL AND ADDITIONAL PROPRIETARY HIGH-LEVEL PROTOCOLS	56 KBIT/S	31	$1,400 TO $14,000	INDUSTRIAL	BUS
DIGITAL COMMUNICATIONS CORPORATION GERMANTOWN, MD.	INFOBUS	OVER 75 MILES	BROADBAND COAXIAL CABLE	CSMA WITH COLLISION DETECTION (CD) AND FDM	1 MBIT/S PER CHANNEL	256	$500 TO $750	HOSPITALS, LARGE CORPORATIONS	BUS (DIFFERENT BANDS)
GOULD INC. (MODICON DIVISION), ANDOVER, MASS.	MODWAY	15,000 FEET	COAXIAL CABLE	TOKEN PASSING	1.544 MBIT/S	250	$1,000 TO $5,000	INDUSTRIAL	BUS
IBM GENERAL SYSTEMS DIVISION ATLANTA, GA.	SERIES/1 RING	5,000 FEET BETWEEN SERIES/1s	COAXIAL CABLE	CSMA/CD VARIATION	2 MBIT/S	16 SERIES/1s	$3,825	DATA PROCESSING	RING
INTERACTIVE SYSTEMS 3M, ANN ARBOR, MICH.	VIDEODATA	40 MILES	BROADBAND COAXIAL CABLE	TIME DIVISION MULTIPLEXING	100 KBIT/S PER CHANNEL	248 PER CHANNEL	$600 TO $900	GENERAL PURPOSE	BUS
LOGICA LTD LONDON, ENGLAND	POLYNET	NA	BASEBAND COAXIAL CABLE	EMPTY SLOT	10 MBIT/S	NA	$1,600	DATA PROCESSING	RING

TABLE 6. LOCAL NETWORKS AT A GLANCE (continued)

During the past several years, the proliferation of local networks has made it difficult to keep track of all the players in the field. Here, a local-networking vendor table sorts out many of the major available offerings.[1]

[1] THIS TABLE IS A REPRESENTATIVE LIST OF AVAILABLE LOCAL NETWORKS AND IS NOT NECESSARILY EXHAUSTIVE

[2] DISTANCES THAT SPAN MILES REQUIRE THE USE OF REPEATERS

[3] THE PRICE PER CONNECTION MAY NOT ALWAYS BE DIRECTLY COMPARABLE BECAUSE OF THE TYPES AND MIX OF PRODUCTS CONNECTED TO THE NETWORK

[4] NOT AVAILABLE

[5] THE DATA RATE IS SIGNIFICANTLY REDUCED WHEN A LARGE NUMBER OF CONNECTIONS ARE MADE

	NETWORK NAME	MAXIMUM DISTANCE SPANNED[2]	TRANSMISSION MEDIUM	ACCESSING SCHEME	SPEED	MAXIMUM NUMBER OF CONNECTIONS	APPROXIMATE PRICE PER CONNECTION	MAJOR APPLICATION	TYPE OF CONFIGURATION
NESTAR SYSTEMS INC PALO ALTO CALIF	CLUSTER/ONE	1,000 FEET	16-WIRE CABLE (FLAT RIBBON, TWISTED-PAIR, OR ROUND SHIELDED BUNDLE)	CSMA/CD	240 KBIT/S	65 APPLE MICRO COMPUTERS	$395	SMALL BUSINESS	SEVERAL TYPES
NETWORK SYSTEMS CORPORATION, MINNEAPOLIS MINN	HYPERCHANNEL (HC) HYPERBUS (HB)	3,000 FEET	COAXIAL CABLE	PRIORITY CSMA	50 MBIT/S (HC) 6.3 MBIT/S (HB)	256 (HC)[5] 256 (HB)	$34,000 (HC) NA (HB)	LARGE MAIN FRAME TO MAINFRAME (HC) OFFICE AND DATA PROCESSING (HB)	BUS
PRIME COMPUTER INC., FRAMINGHAM MASS	PRIMENET	750 FEET BETWEEN NODES	COAXIAL CABLE	TOKEN PASSING	8 MBIT/S	OVER 200	$9,000	GENERAL PURPOSE	RING
STRATUS COMPUTER INC NATICK MASS	STRATALINK	750 FEET OR 1,500-FOOT INTERVALS WITH REPEATERS	BASEBAND COAXIAL CABLE	TOKEN PASSING	2.8 MBIT/S	32 STRATUS COMPUTERS	$5,000 PER MACHINE, WHICH SUPPORTS 64 TERMINALS	INDUSTRIAL HOSPITALS AND AIRLINES	RING
SYTEK INC SUNNYVALE, CALIF	LOCALNET (SYSTEM 20 AND SYSTEM 40)	5 TO 20 MILES	BROADBAND COAXIAL CABLE	CSMA/CD	128 KBIT/S TO 2 MBIT/S	16,000 TO 64,000	$500	GENERAL PURPOSE	BUS
UNGERMANN BASS INC SANTA CLARA CALIF	NET/ONE	4,000 FEET	BASEBAND COAXIAL CABLE	CSMA/CD	4 OR 10 MBIT/S	200	$500 TO $1,000	OFFICE AND DATA PROCESSING	BUS
WANG LABORATORIES INC., LOWELL, MASS	WANGNET	2 MILES	BROADBAND COAXIAL CABLE	CSMA/CD AND FDM	9.6 KBIT/S TO 12 MBIT/S	512 TO SEVERAL THOUSAND DEPENDING ON BAND AND EQUIPMENT	$400 TO $1,500	GENERAL PURPOSE	BUS
XEROX CORPORATION EL SEGUNDO CALIF	ETHERNET	1.5 MILES	BASEBAND COAXIAL CABLE	CSMA/CD	10 MBIT/S	100 TO 1,000 DEPENDING ON TRAFFIC	$1,200	OFFICE	BUS
ZEDA COMPUTERS INTERNATIONAL LTD PROVO UTAH	INFINET	3.1 MILES	TWISTED PAIR WIRE	CSMA/CD	25 KBIT/S	30 MICRO COMPUTERS	$100	SMALL BUSINESS	BUS
ZILOG INC CUPERTINO CALIF	Z NET	1.2 MILES	BASEBAND COAXIAL CABLE	CSMA/CD	800 KBIT/S	255	NA	OFFICE AND DATA PROCESSING	BUS

The specific needs of different users have been examined in a recent analysis of LANs (Venture Development Corporation, 1981). This analysis suggests that for large data processing systems, including larger LANs that interconnect smaller ones, and for real time process control, TDMA (time division multiple access or TDM) will be used. Token passing will be preferred for factory automation systems and laboratory data collection systems. For office automation systems where traffic is infrequent, but intense when needed, the CSMA/CD access method will be more appropriate. These conclusions are very general and may not apply to given users in any of these application areas. Cost and availability of commercial LANs may be as big a factor as any other consideration.

One of the major factors involved in the decision to use a broadband LAN is whether there will be any need for video or image transmission. Video data is relatively common in many operations in terms of security monitoring, local training functions, and video conferencing. Many people predict a growing use of video conferencing as a way to reduce travel costs for various business and technical meetings. If video conferencing is an important factor in the future plans of a user, then a broadband LAN should be given serious consideration. If a long-haul video network is already in operation by the user, then it is very likely that the LAN should be connected to this system, and should offer video capability. The use of image data transmission is also of importance for many users, where data such as that from medical CAT scanners or industrial CAD/CAM systems are involved. Applications such as these are increasing at a rapid rate, and this possibility should be carefully examined by potential users.

INTERCONNECTION OF NETWORKS

Once an LAN has been established, it is often desired to interconnect that LAN with other nearby LANs, with other long-haul networks or with other networks offering general data services. A large office building, a large industrial plant or a large military base might all have a collection of various LANs at their facility. Interconnection of these various LANs may offer capabilities to share scarce resources, improve system reliability, and transmit information to different corporate or command levels. For a new installation, a broadband cable network with a number of separate channels for different LANs is possible, and often offers the easiest and cheapest way to interconnect them. Otherwise, specialized interface communication processors, or gateways, must be added to each LAN to allow an interconnection.

One of the most popular gateways for any LAN is an X.25 gateway to the public packet switching network. For military applications, a gateway to the DOD ARPANET might be very desirable. Many of the commercial vendors for LANs recognize the need for these interconnections and are working on supplying the appropriate gateways. These network interconnections may also provide a means for different types of computers and computer systems to directly communicate with each other. Existence of these gateways should be a substantial consideration when examining commercial LANs for possible purchase and application.

LOCAL AREA NETWORK STANDARDS

Many different LANs are offered by a variety of vendors, and they are not standardized to any great degree. An exception is Ethernet, which is standardized among Xerox, Intel Corporation, and Digital Equipment Corporation. The companies offering LANs recognize the need for standardization, and several industry committee groups are trying to arrive at an acceptable standard. The Institute of Electrical and Electronic Engineers (IEEE) has a committee working on a LAN standard known as IEEE 802. This IEEE standard has been evolving in a manner that attempts to accommodate many diverse application areas and functional requirements. The standardization efforts are associated with specifying the protocols used on LANs, and so far the IEEE 802 committee has been concentrating on the level 1 and 2 (physical and electrical interface) layers of the protocol. Many problems remain at the higher protocol levels, and the IEEE 802 standard may turn out to allow many different alternatives for data rates, error detection codes, and access methods. In the meantime, while the IEEE 802 committee is slowly making progress, more vendors are turning out their own LAN systems.

LOW COST COMPUTER NETWORKS

The widespread availability of inexpensive microcomputers such as the Apple II, Radio Shack TRS-80 and the Atari 400 and 800 has resulted in many organizations purchasing a number of these machines for educational, business and training functions. Schools and universities have been eager to purchase large numbers of these inexpensive computers to use in teaching and research. Many teaching laboratories have been established to use groups of 10-50 microcomputers for various learning and training activities. Usually these applications are not sufficiently funded in the sense that not every microcomputer is provided with a set of peripherals such as one or more floppy disk memory units, a hard disk unit, and a reasonable speed printer. While several hard disks and high speed printers may be available, they are only connected to a few microcomputer systems.

Several companies have means of connecting a group of microcomputers so each one can access a disk unit or a printer connected to only one of the microcomputers. The Corvus Omninet and the Nestar Systems Cluster/One are the most sophisticated of these LANs specifically for interconnecting up to 64 microcomputers. Omninet is a 1 megabit per second network using a single twisted pair wire while Cluster/One has a 240 kilobit per second rate using a 16 wire ribbon cable. A lower speed network is INFINET produced by Zeda Computers, International which connects up to 30 microcomputers with a data rate of 25 kilobits per second. All of these LANs are baseband mode systems and are limited in their general utility. However, for a facility that already has a group of microcomputers, these networks may be cost effective considerations for specific applications.

There are also lower cost but less capable interconnection networks for certain microcomputers than those described above. Corvus makes a version of its Omninet that is called Appleshare which is designed to allow up to 63 Apple II computers to share one to six diskette drives and a common printer.

Low cost control cards and cables are also available for both Apple II and Radio Shack TRS-80 microcomputers to allow many microcomputers to share a common disk drive unit. With these lower cost approaches comes a reduction in the capability of service, but still they may have desirable features to offer for their price.

NAVY COMPUTER BASED INSTRUCTION APPLICATIONS

There are many possibilities for applications of LANs to instructional computer systems in general, and to Navy CBI systems in particular. Numerous classroom applications can be envisioned where a group of 10 to 30 microcomputer-based individualized instruction stations are involved. These instruction stations can be interconnected with a LAN that is self-contained for that classroom, or can be interconnected with a LAN that covers a single building with several school rooms. Again, this LAN can be restricted to a single building, or it could be part of a larger LAN covering several buildings or an entire multibuilding facility.

The design possibilities for using LANs are quite numerous as the above discussion implies. Layers of LANs can be used, or a single LAN can be used in many applications. Which approach is the best is probably dependent on the overall system goals and the costs of the various LANs. The use of a broadband LAN with FDM/TDM providing a series of separate channels may be a reasonable possibility. The general consideration for baseband or broadband mode is determined by digital data speed requirements and whether there is a need for audio data, video conferencing, or image data transmission. The use of video conferencing should always be a choice examined in an instructional system network design. It may be discarded as an unnecessary capability in many instances, but it should always be considered. The use of image data for certain training needs may also be a consideration. High resolution, digitally coded images are presently being used for industrial CAD/CAM applications, and these same techniques may offer some potential benefits for military training systems.

This subsection has briefly touched on several areas where LANs can be applied successfully to Navy CBI systems. However, this is only part of the overall task of using distributed processing systems in large scale Navy CBI systems. The next section of this report is devoted to discussing the application of the distributed system concepts presented up to now to typical large scale systems for computerized instruction in the Navy.

SECTION VII

DISTRIBUTED PROCESSING AND COMPUTER BASED INSTRUCTION

This section integrates the concepts and design consideration introduced in previous sections through discussion of their application to Navy computer-based instructional systems. For the purpose of this discussion it will be assumed that individualized instruction will be widely utilized, and a large scale computer/communications system will be implemented as the overall system design. The former assumption is not necessary, however, to achieve benefits from the use of stand-alone microcomputers for instruction in the NAVEDTRACOM.

A GENERIC CBI SYSTEM

In order to establish CBI system design goals, and to evaluate alternative system configurations, it is necessary to identify the functions the system will be expected to perform in a typical environment. Once these functions are identified, alternatives can be discussed in terms of their effectiveness and efficiency in meeting system requirements.

Five categories of computer-based instruction findings are shown in figure 13. Computer-Based Instructional Administrative Support includes record keeping, scheduling, reporting, resource management, and other instructional functions required for effective program administration. Computer-Based Instructional Delivery includes all computer-based media options such as classical computer-aided instruction (CAI), computer graphics, computer controlled video, computer simulation, and interactive computer-based 2-D and 3-D applications. Computer-Based Instructional Testing includes test generation, storage, administration, and evaluation. Computer-based Instructional Management includes those functions which guide a student through an instructional sequence such as test response analysis, prescriptive guidance, and student monitoring. Computer-Based Curriculum Development includes computer-aided course authoring with both text and graphic formats and may include components of the Navy Technical Information Presentation Program (NTIPP) or the Computer Readability Editing System (CRES).

With this general definition of CBI system functions, it is possible to develop a specification for a generic Navy training CBI system. The generic application is envisioned to include a classroom level with individualized instruction student stations; a schoolhouse level with coordination and control functions for a number of classrooms; a training center level with coordination and control functions for a number of schoolhouses and an instructional program development function; and a top administrative level which exercises system control and management. The generic CBI system is shown in figure 14. A more detailed structure of the classroom level is shown in figure 15. Each classroom is assumed to consist of a number of individualized instruction student work stations connected to a control computer with a hard disk storage unit.

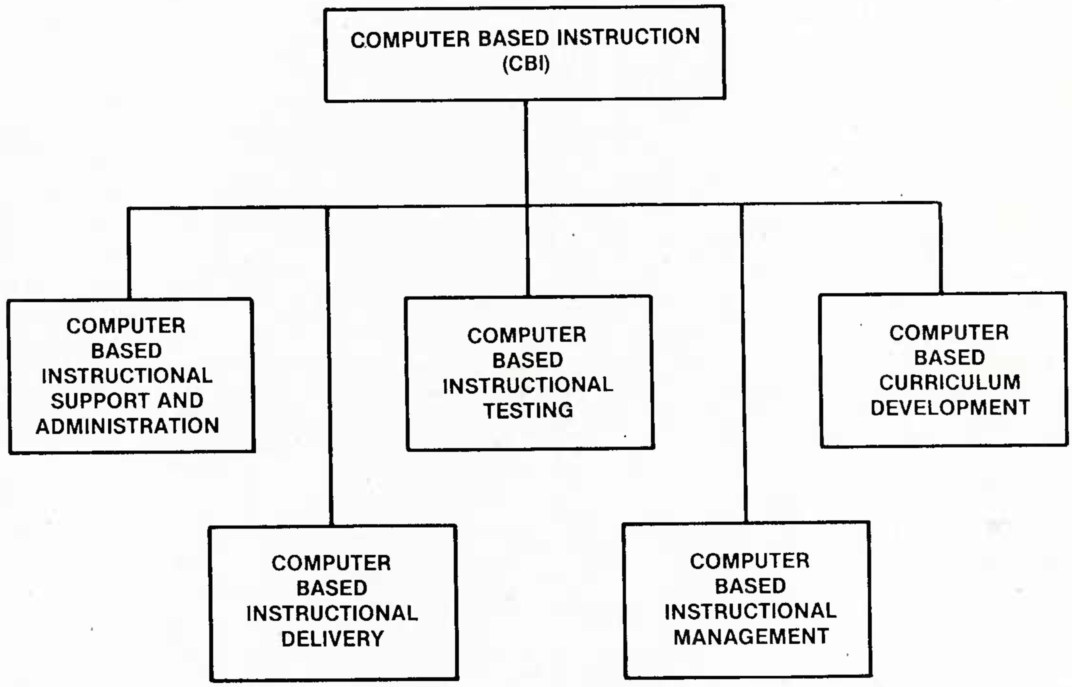


Figure 13. Major Categories of CBI

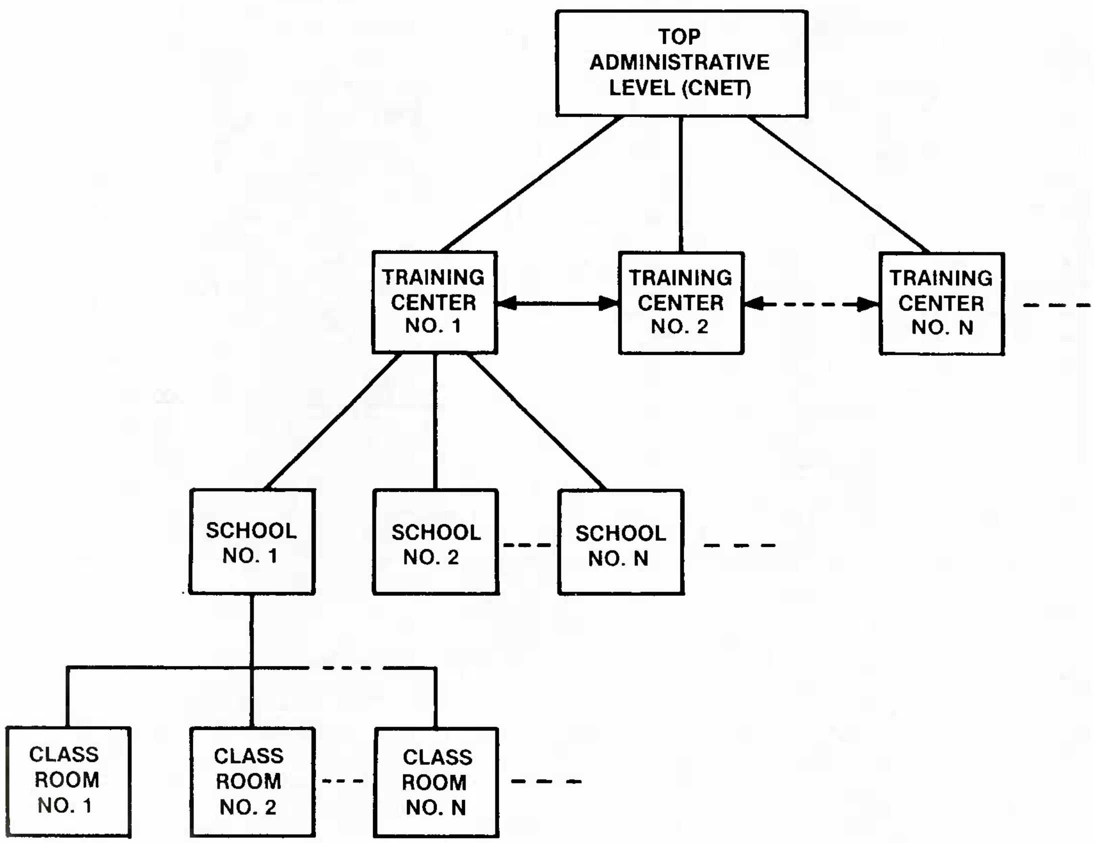


Figure 14. Generic Structure of a Navy Distributed CBI System

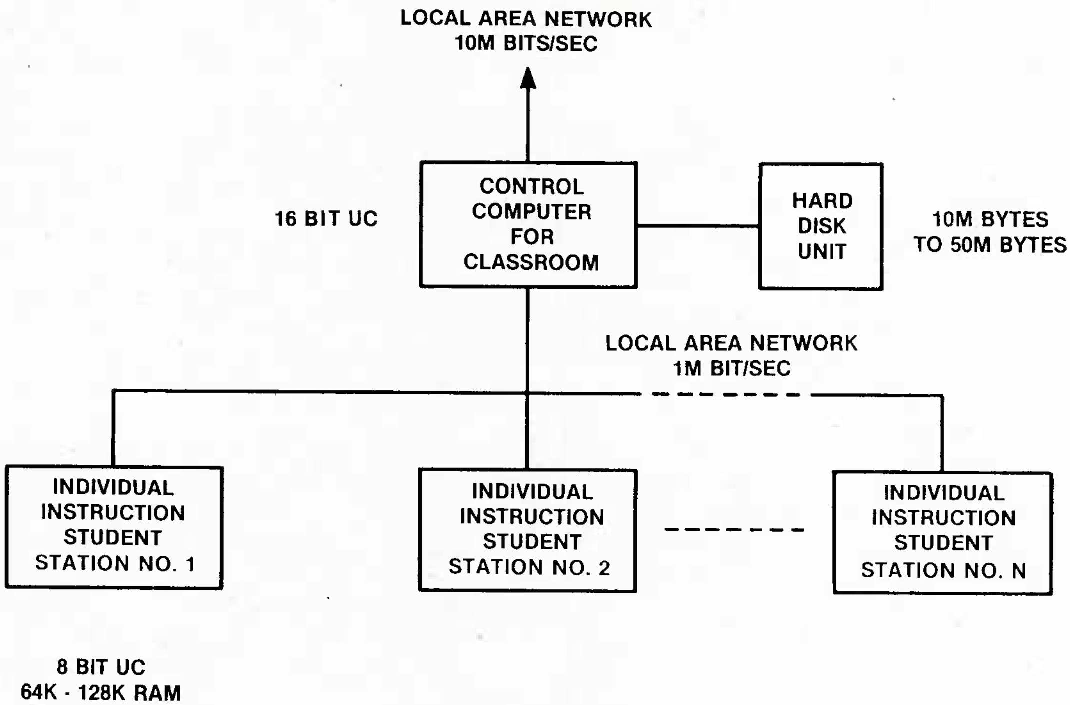


Figure 15. Classroom Configuration for Distributed CBI System

DESIGN FEATURES

The overall structure of this generic CBI system is hierarchical since it must integrate with the Navy training system organization. The lowest level in the hierarchy is the classroom, with a number of individual instruction student stations. These student stations are, in essence, intelligent terminals for the delivery of instruction and for the collection of student performance data. It is also assumed that these student stations are connected together in a local area network. Each of the training centers is considered to be within a single geographical location such as a base, such that all of the classrooms in a school and all of the schools in a training center are interconnected with a local area network. An Instructional Program Development Center (IPDC) is also assumed to be interconnected to the local area network associated with a training center. A high speed communication interconnection between training centers will be required to satisfy high capacity data transmission requirement. An irregular mesh interconnection will allow training centers and the top administrative level to share data.

More detailed specifications of a CBI system require the availability of information regarding the quantities of data needed for effective system operation. As an example, assume a classroom or group of classrooms supporting a single course and a school supporting a maximum of 10 courses. Thus, with 40 classrooms per school there will typically be four classrooms devoted to the same course. Each course is assumed to consist of 20 modules, with each module representing 16 hours of CBI system instruction. For the purpose of this example, modules consist of eight lesson topics of 2 hours duration each, and a lesson topic is divided into four segments of one-half hour instruction each. The actual content of material in a segment, as it relates to system computer speed and capacities, depends on the nature of the CBI system method of instruction. If classical CAI is involved, each segment is assumed to require 12 display screens of text and six display screens of graphics. If interactive graphics and simulation are involved, as in Performance-Based Instruction (PBI), then fewer pages of text and more pages of higher resolution graphics are required.

To estimate the size and capability of the microcomputers needed to support CBI functions at the classroom level, an assumption must be made relating the quantity of text and graphics and the quantity of random access memory (RAM). With an estimate of 2K bytes of RAM memory for each screen of text and 4K bytes of RAM for each screen of conventional graphics, a total RAM requirement of 48K is necessary for each segment--or one-half hour of instruction. If an additional 16K bytes are allotted for the various control and overhead programs needed to operate the entire system, a total of 64K RAM will be necessary to store the information needed to provide a student with a nominal half hour of conventional CAI instruction. This planning factor may be an underestimate; therefore, a range of 64K to 128K is suggested as a planning factor to provide an adequate margin for design contingencies. With the more sophisticated interactive graphics and simulation required for PBI, this requirement could grow to 256K bytes. These estimates of computer memory for a segment (one-half hour) of CBI instruction form the basis for data transfer rates needed for the LANs and the size of disk storage at various locations within the system.

The amount of disk memory needed to accommodate an entire course is easily computed using the previously determined storage requirements of 64K-128K bytes for each course segment. One course contains 20 X 8 X 4 = 640 segments, or 640 X 64K to 640 X 128K bytes, for a total of 41M - 82M bytes of data. There are many alternatives for distributing this data in the system. Each classroom could be provided with a 50M - 90M disk to contain all of the course data. However, with this arrangement, each classroom disk would duplicate the data files. As an alternate arrangement, four 10M - 20M disks could be used at the classroom level to distribute the data base into non-replicated pieces or four 20M - 40M disks could be used to store one-half the total course file each, thereby providing backup for failed disks. As another option, one 50M - 100M disk could be provided at the school level to be accessed by each of the classrooms supporting that one course. These are only some of the numerous possibilities for the data file configuration for an actual distributed processing system.

The individual student stations are assumed to be interconnected through a local network. A star, ring/loop or a multidrop bus based LAN could be used. Since many LAN analysts suggest that bus structures are generally the most desirable, this may be the preferred configuration. The required data rate can be determined by examining the time needed to transfer the 64K-128K of data to the student stations. The following table shows the times to transmit 64K-128K bytes (1 byte = 8 bits) at various data rates.

TABLE 7. DATA TRANSMISSION TIME

9600 Bits/Sec	54-108 Sec
19.2K Bits/Sec	27-54 Sec
240K Bits/Sec	2.2-4.4 Sec
1M Bits/Sec	0.5-1.0 Sec
10M Bits/Sec	0.05-0.1 Sec

Taken individually, these transmission times--or delay times--might be acceptable. However, if many of the 20 student stations request data files at the same time, the total delay time for a given student could be critical. The 240K bit/sec rate is considered to be the lowest acceptable value, since this would give a 22-44 second delay if 10 student stations initiated requests at the same start-up time. The 1M bit/sec rate is better and, if the cost is not significantly different, it should be selected. This data rate reduces the system delay time to 5-10 seconds if 10 students initiated simultaneous requests. The 10M bit/sec rate is the best, of course, but this probably represents over-capacity that will not likely be cost-effective. The Omninet LAN from Corvus Systems, Incorporated is a 1M bit/sec network that will interconnect up to 64 microcomputers at a cost of $500-$1,000 per computer. This appears to be a cost-effective configuration.

The student stations should be microcomputer based, and must have memory capacities of 64K-128K of RAM. For conventional CAI instruction, the speed of current 8 bit microcomputers is probably sufficient to satisfy this requirement. In the near future it is expected that16 bit microcomputers with 128K-256K of RAM will be available for $1,000 each. The higher speed

16 bit microcomputer student stations will be required if PBI involves the display of interactive graphics and 2-D and 3-D simulation. Most 16 bit microcomputers will address well over 256K of RAM, so this is not a problem as it is with 8 bit microcomputers. Thus, it is likely that 16 bit microcomputers will represent the best choice for future applications of this type.

The classroom control computer must interconnect with the classroom LAN plus connect to a disk storage unit of 10M byte-80M byte depending on how the files are distributed and to what extent PBI is used. Since this data would be addressed and transferred in blocks of 64K-128K-256K, a 16 bit microcomputer could still handle this task. A typical microcomputer to use at this level is the Corvus Systems, Incorporated Concept computer which presently costs about $4,000. This computer is designed to interface with the Corvus Omninet which would satisfy the classroom level network requirements.

Schools form the next level in the hierarchical system, and support 40 classrooms and 10 courses each. This level supports and coordinates the functions of the classrooms, and transmits management information to the higher levels in the system. A minicomputer such as a WANG VS might be a typical selection at this level depending on how much of the data base was located at this level, and a 400M-800M disk would be needed. If the course data base is distributed throughout the classroom level, then there would not be a need for this size computer or disk at the school level. The data rates into and out of the school level must be higher than those in the individual classroom level because of the necessity to transfer entire course files from the training center with its IPDC function. A data rate of 10M bit/sec would probably be needed to interconnect each school both up and down the system.

Since the training center is considered to be within one geographic location, the same LAN can be used to connect the classrooms to the schools, the schools to their training center, and the IPDC to the training center. A 10M bit/sec LAN could be used for this total task. This LAN could be baseband or broadband to handle the 10M bit/sec rate. Ethernet would be a possible choice for a baseband network. However, with the view that video transmission, video conferencing, and image data might be desirable either at the time of system installation or in the future, it would seem prudent to use a broadband network. If the network will only be used for digital data at the 10M bit/sec rate, then low cost, single channel RF modems can be used. This should cost no more than a 10M bit/sec baseband network. At some time in the future, when additional services will be needed or desired, the greater capacity of the broadband LAN could be utilized by adding multichannel RF modems.

The training center would be the highest technical information level in the system. It would (1) serve as the repository for the entire system data files, (2) serve as the location for new course development and for course revision, and (3) be the technical control center for the lower levels of the system that it served. This level would be connected to the 10M bit/sec LAN which ties it to the schools and the IPDC. The training center would probably need several gigabytes of disk storage (1 gigabyte equals 1,000 megabytes) to accommodate the master course files on the system. However, this could again be distributed and backed up by replicating the files at other locations. A WANG VS or similar capacity system would be needed at this level.

The communication needs for the training centers include a high speed link to the other training centers and to the top administrative level (CNET). Because of the need to transfer large blocks of data--such as complete course files--between training centers, there must be access to a high speed data communications line. However, this high speed link would probably not be used continuously. Thus, a reasonable approach to this need would be to use the public packet switching network. This would allow high speed transmission at the infrequent times it was necessary, and the cost would only be paid for those actual transmission times and quantities of data involved. In addition to this high speed, low frequency of use transmission link, the training centers need lower speed, fairly continuous use transmission links to the top administrative level. These links allow the flow of summarized management information from the lower levels of the system to the top where the major operational decisions are made. Dedicated 1200 baud communication channels from each training center to the top administrative level would probably fulfill the anticipated data exchanges.

The IPDC function at each training center requires special attention because of the need to provide facilities for course authoring of both text and graphics materials. Special purpose authoring work stations will be needed to provide text processing, editing and formatting capabilities as well as graphics production and merging capabilities. Several such graphics work stations are presently available, and are based on 16 bit microcomputers. These include the SUN work station, the Xerox 1100 Smalltalk-80, and the Fortune 32:16 system. Prices range from $10,000 to $20,000 each for these high resolution, high capability text and graphics processors. These special purpose authoring work stations would then be interconnected to the 10M bit/sec LAN at the training center. Thus, easy access to large disks and other system resources would be available. In addition, the authoring work stations would be able to download new or revised courses to the schools; to be further downloaded to classrooms for instructional delivery.

Finally, at the top administrative level, there will be a need for a computer with a large disk storage capacity to gather, process, and store the data necessary to maintain management control of the entire CBI system.

The size of this facility is dependent on the quantity of information needed for effective management. No general recommendations can be made for this level until all of the information needs are clearly specified.

CONCLUSION

In this section, an attempt has been made to integrate and synthesize the basic concepts and design considerations of distributed systems by examining an example of a Navy CBI system. This example application has been far from complete in the sense that a simplified CBI system model has been discussed, and only a broad examination of this simplified system has been employed. While this discussion has been severely restricted, it provides a review and synthesis of the many detailed topics presented in the previous section. All of the topics addressed are interrelated through the system design, and each must be considerd in light of this dependency. Although an example with greater specificity would be more informative, it would require a lengthy and detailed discussion which is beyond the scope of this report. Consequently, this section has been structured to offer additional insight to the reader without adding unnecessary complexity.

REFERENCES

Booth, G. M. The Distributed System Environment. New York: McGraw-Hill, 1981.

CNET Instruction 3920.1B, Research, Development, Test and Evaluation Program of the Naval Education and Training Command, August 26, 1981. Chief of Naval Education and Training, Naval Air Station, Pensacola, FL 32708.

CNET Instruction 5260.1, Computer Managed Instruction (CMI) System, February 17, 1977. Chief of Naval Education and Training, Naval Air Station, Pensacola, FL 32708.

CNTECHTRA Instruction 5400.7A, Computer Managed Instruction, 22 May 1974. Chief of Naval Technical Training, Naval Air Station, Memphis (75), Millington, TN 38054.

Data Communications. "Local Networks at a Glance." Data Communications. December 1981.

Freeman, H. A. and Thurber, K. J. Tutorial on Microcomputer Networks. Los Angelos, CA: IEEE Computer Society Press. 1981.

Gibson, R. W. "Evaluating Your Local Network Needs." Small Systems World, 1982, 10(5), 25-27.

Katzan, H. Distributed Data Processing. New York: Petrocelli Books, 1978.

Katzan, H. Distributed Information Systems. New York: Petrocelli Books, 1979.

Krutsch, T. E. "A User Speaks Out: Broadband or Baseband for Local Nets?" Data Communications, December 1981.

Levy, W. A. and Mehl, H. "Taking a New Look at Matrix-Switched Systems." Mini-Micro Systems, March 1982.

McGlynn, D. R. Distributed Processing and Data Communications. New York: Wiley-Interscience Press, John Wiley and Sons, 1978.

Micheli, G. S., Morris, C. L., and Swope, W. M. Computer Based Instructional Systems -- 1985 to 1995. TAEG Report No. 89, August 1980. Training Analysis and Evaluation Group, Orlando, FL 32813 (ADA088702).

Middleton, M. G., Papetti, C. J., and Micheli, G. S. Computer Managed Instruction in Navy Training. TAEG Report No. 14, March 1974. Training Analysis and Evaluation Group, Orlando, FL 32813 (AD777040).

Sherman, K. "RS-449 - The New Interface Specification." Small Systems World, March 1982.

REFERENCES (continued)

Small Systems World. "Getting a Line on Data Communications." 1981, 8(8), 21.

Small Systems World. "Local Area Networks to Triple by 1986." Small Systems World, March 1982.

Tektronix, Inc. Essentials of Data Communication. 1978. Tektronix, Inc., Beaverton, OR.

Venture Development Corporation. Local Area Networks: A Strategic Analysis, 1981-1990. 1981. Wellesley, MA.

Weitzman, C. Distributed Micro-Minicomputer Systems. Englewood Cliffs, NJ: Prentice-Hall, 1980.

Zajkowski, M. M., Heidt, E. A., Corey, J. M., Mew, D. V., and Micheli, G. S. An Assessment of Individualized Instruction in Navy Technical Training. TAEG Report No. 78, November 1979. Training Analysis and Evaluation Group, Orlando, FL 32813 (AD A080183).

DISTRIBUTION LIST

Navy

OASN (M&RA)
CNO (OP-115, OP-987H, OP-987, OP-11, OP-12)
ONR (442 (3 copies), 270)
CNM (MAT-0722)
CNET (01, 02, N-5)
CNAVRES (02)
COMNAVAIRSYSCOM (03, 340F, 413E)
CNTECHTRA (016 (5 copies), N-6)
CNATRA (Library (2 copies))
COMTRALANT (00)
COMTRALANT (2 copies)
COMTRALANT (Educational Advisor)
COMTRAPAC (2 copies)
CO NAVPERSRANDCEN (Library (4 copies))
NAVPERSRANDCEN Liaison (021)
Superintendent NAVPGSCOL (2124, 32)
Superintendent Naval Academy Annapolis (Chairman, Behavioral Science Dept.)
CO NAVEDTRAPRODEVCEN (Technical Library (2 copies), PDM)
CO NAVEDTRASUPPCENLANT (N-3 (2 copies))
CO NAVEDTRASUPPCENPAC (5 copies)
CO NAVAEROMEDRSCHLAB (Chief Aviation Psych. Div.)
CO FLECOMBATRACENPAC
CO FLECOMBATRACENLANT
CO NAMTRAGRU
CO NAVTECHTRACEN Corry Station (101B, 3330, Cryptologic Training Department)
CO NAVTRAEQUIPCEN (TIC, N-001, N-002, N-09P)
Center for Naval Analyses (2 copies)
OIC NODAC (2)
CO TRITRAFAC (2 copies)
CO NAVSUBTRACENPAC (2 copies)
CO FLEASWTRACENPAC
CO FLEASWTRACENLANT
CO NAVSUBSCOL NLON (Code 0110)
CO NAVTECHTRACEN Treasure Island (Technical Library)
TAEG Liaison, CNET 022 (2 copies)
DIR NAVEDTRAPRODEVCENDET Memphis
CO NAVTECHTRACEN Meridian
COMFLETRAGRU Pearl Harbor
DIR NAVEDTRAPRODEVCENDET Meridian
CNET Liaison Officer, Williams Air Force Base
DIR NAVEDTRAPRODEVCENDET GLAKES
CISO, SERVSCOLCOM GLAKES
CISO, NTTC Meridian
CO NAVAEROSPMEDINST (Code 13, Code 11)
CO FLETRACEN, Mayport
CO FLETRACEN, San Diego
CO FLETRACEN, Norfolk
CO FLEMINEWARTRACEN
CO HSL 31 (Training Dept.)

DISTRIBUTION LIST (continued)

Air Force

Headquarters, Air Training Command (XPTD, XPTIA) Randolph Air Force Base
Air Force Human Resources Laboratory, Brooks Air Force Base (2 copies)
Air Force Human Resources Laboratory (Library), Lowry Air Force Base
Air Force Office of Scientific Research/NL
Headquarters Tactical Air Command (DOOS), Langley Air Force Base
AFMTC/XR, Lackland Air Force Base
Headquarters 34 TATG/IDM, Little Rock Air Force Base
Headquarters MAC/DOT, Scott Air Force Base
4235 Strategic Training Squadron, Carswell Air Force Base

Army

Commandant, TRADOC (Technical Library)
ARI (Technical Director, PERI-RH, PERI-SM, PERI-IC, Library (2 copies))
ARI Field Unit - Fort Leavenworth
ARI (Reference Service)
ARI Field Unit - Fort Knox (PERI-IK)
COM USA Armament Materiel Readiness Command (DRSAR-MAS)
COMDT, USAIPRM (ATSG-DT-R)
USA Missile & Munitions Center and School (ATSK-AIT)

Coast Guard

Commandant, Coast Guard Headquarters (G-P-1/2/42, GRT/54)

Marine Corps

CMC (OT)
CGMCDEC
Director, Marine Corps Institute
CO MARCORCOMMELECSCOL

Other

Military Assistant for Human Resources, OUSDR&E, Pentagon
Institute for Defense Analyses
COM National Cryptologic School (Code E-2)
Old Dominion University

Information Exchanges

DTIC (12 copies)
DLSIE
Executive Editor, Psychological Abstracts, American Psychological Association
ERIC Processing and Reference Facility, Bethesda, MD (2 copies)